PM-67-21

NASA TMX 60900

SECOND COMPILATION OF PAPERS ON TRAJECTORY ANALYSIS AND GUIDANCE THEORY

NASA

ELECTRONICS RESEARCH CENTER
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

PM-67-21

SECOND COMPILATION OF PAPERS ON TRAJECTORY ANALYSIS AND GUIDANCE THEORY

JANUARY 1968

Prepared by Contractors
for the Guidance Laboratory
NASA Electronics Research Center
Cambridge, Massachusetts

FOREWORD

Compiled in this volume are twelve papers from agencies working with the Guidance Laboratory of NASA-ERC. These papers concern special studies in the disciplines of guidance theory, trajectory analysis, and celestial mechanics. They include:

1. A presentation of a method for classifying and analyzing guidance modes, together with a survey and classification of existing modes;

2. A development of a minimum fuel guidance solution for certain hyperbola-circle transfers;

3. A presentation of generalized necessary conditions for a minimizing control function;

4. A demonstration of how many non-standard problems in control theory can be transformed to "ordinary differential form" and treated in a standard way;

5. A study of impulsive transfer optimization, presenting a computational scheme for determining optimum n-impulse trajectories;

6. A demonstration of the utility of a group theoretical principle in solving non-homogeneous linear systems of differential equations and an application to the perturbation of elliptic motion;

7. A derivation of a useful form of the Sundman inequality;

8. A study of a formal approximate decoupling of the n-body problem into a k-body problem ($k < n$) and an $(n - k + 1)$ - body problem;

9. A presentation of theory for the solution of the perturbation problem for a system of non-linear differential equations under general linear two-point boundary conditions;

10. A presentation of an algorithm for automatic computation of derivatives of a composite function and its application to a power series solution of a celestial mechanics problem;

11. A presentation of results on the long-term behavior of close lunar orbiters;

12. An analytical study of the three-dimensional stability of motion of a particle near L_4 in a non-linear Earth-moon gravitational field.

The first paper provides a framework enabling many specific studies to result in an accumulating body of guidance mode information, useful to both the guidance theorist in research work, and the guidance software engineer in planning and development work.

The second paper provides one guidance solution of interest to the guidance theorist, and could form the basis for a guidance mode to be studied from the viewpoint of the first paper.

The third paper provides the trajectory analyst with tools for attacking a broader class of optimization problems, and the fourth paper points out a way he may profitably apply a large body of results to his problems which are not of the "ordinary differential" type.

The fifth paper aids in the search for optimum impulsive orbit transfers, and the sixth through the ninth papers contribute celestial mechanics theory useful toward the solution of problems occurring in mission design, orbit determination or orbit prediction.

The tenth paper supports both automatic symbolic processing efforts and computational requirements in trajectory analysis and celestial mechanics.

The last two papers aid the trajectory analyst or mission designer interested in stable orbits near the moon or its equilateral libration points.

SUMMARY

This volume contains technical papers on NASA-sponsored studies in the areas of trajectory analysis and guidance theory. These papers cover the period beginning 1 October 1966 and ending 1 October 1967. The technical supervision of this work is under the personnel of the Guidance Laboratory at NASA-ERC.

CONTENTS

INTRODUCTION

This document contains twelve technical papers covering work sponsored by the Guidance Laboratory of the NASA Electronics Research Center (ERC). The papers are concerned with guidance theory, trajectory analysis, and celestial mechanics.

The following table presents the contributing institutions and the discipline of the paper.

Institution/ Company	Author	Discipline
TRW	C.G. Pfeiffer	Guidance Theory
AMA	T.N. Edelbaum	Guidance Theory
Northeastern Univ.	J. Warga	Trajectory Analysis*
Princeton Univ.	P.M. Lion	Trajectory Analysis
CRA	D. Lewis/P. Mendelson**	Celestial Mechanics
IBM	P.Sconzo/D. Valenzuela	Celestial Mechanics
Stanford Univ.	J. Vagners/H.B. Schechter	Celestial Mechanics

* Two papers

** Four papers

Synopses of the individual papers are presented below:

<u>Paper No. 1</u>

The first paper by C.G. Pfeiffer of TRW surveys the guidance modes presently in use or in development, with the objective of appraising their usefullness when applied to a unified guidance concept. Unified guidance refers to the application of one hardware system as well as one overall computational scheme for guidance in all phases of a mission.

The breadth of the study called for, and the importance of the consequences drawn from it, made it advisable to review critically concepts and terms in use in space guidance and to update definitions where necessary. Thus this paper starts with defining terms such as guidance, guidance mode, unified guidance and others. The paper also elaborates on the relationship of guidance to navigation.

Chapter III then gives a rather comprehensive survey of presently known guidance modes (USA-developed) and categorizes them according to various principles of division.

For the evaluation of guidance modes in general, as well as with respect to the applicability to unified guidance, measures of performance are suggested in the following chapter. These concern such criteria as optimality, accuracy, applicability, flexibility, and others.

The paper does not aim at evaluating all the listed modes, but rather tries to lay the groundwork for such an attempt.

Paper No. 2

The second paper, a technical progress report from Analytical Mechanics Associates, Inc., by T. N. Edelbaum, presents a minimum fuel guidance procedure. It presents a first-order impulsive correction procedure for mid-course guidance, along with a procedure for establishing approximate equations of motion which are then integrated. This guidance procedure is classified in the paper by C. Pfeiffer (No. 1 in this document).

Paper No. 3

The third paper, written by Jack Warga of Northeastern University, adds directly to the results presented by him in the first compilation in this series.* In the referenced paper, Prof. Warga presented existence and approximation theorems for minimizing controls applicable to a general class of optimization problems. The present paper contains a statement and proof of some corresponding conditions that are necessary for the solutions to be minimizing.

* NASA, First Compilation of Papers on Trajectory Analysis and Guidance Theory, NASA Scientific and Technical Information Division, Wash., D.C., 1967.

INTRODUCTION

Paper No. 4

In the fourth paper, Prof. Warga has presented some observations that appear to have been overlooked by some workers in mathematical control theory faced with problems that are not of the "ordinary differential" type. He illustrates with three specific examples of how one may transform the problem to one which is immediately treatable with classical results of control theory.

Paper No. 5

The fifth technical paper is concerned with trajectory analysis. Written by P.M. Lion of Princeton University, it presents a study of impulsive trajectory calculations. It is based on Lawden's primer vector and presents necessary conditions for a trajectory to be optimum. This primer vector has significance for non-optimal trajectories. When these ideas are combined, a computational scheme for determining optimum n-impulse trajectories is suggested.

The remaining papers are on various aspects of celestial mechanics.

Paper No. 6

The sixth technical paper, written by D.C. Lewis and Pinchas Mendelson, formerly of Control Research Associates and now of Zetesis Corp., exploits a principle which considers a system of differential equations invariant under continuous and differentiable group transformation. It shows that it is possible to write down a number of linearly independent solutions of the variational equations equal to the number of independent parameters of the group. This exploitation is used to present several solutions to the Keplerian case.

Paper No. 7

The seventh technical paper, written by D.C. Lewis, and entitled, "Comments on the Sundman Inequality," presents some preliminary theorems in vector analysis and applies them in developing the Sundman inequality in the form used in former papers.

Paper No. 8

The eighth paper, also written by D.C. Lewis, discusses partial and complete decoupling of a n-body problem into a k-body problem ($k < n$) by use of developed parameters of the problem. Theory is presented to develop the so-called quasi-first integrals of the partially decoupled system from the first integrals of the unreduced system.

Paper No. 9

The ninth paper, also written by D.C. Lewis, discusses the perturbation problem for a system of non-linear differential equations. The problem is reduced to solving k "bifurcation" equations where k is the degeneracy of the problem. A Green's matrix for linear systems with boundary conditions of arbitrary degeneracy is constructed.

Paper No. 10

The tenth paper, written by P. Sconzo and D. Valenzuela of IBM, presents work done on automatically computing the derivatives of a function by the recursive Schlömilch-Cesàro formulation. The general expression obtained was used to construct the power series expansions in the time variable of those powers of the radius vector which appears most frequently in celestial mechanics.

Paper No. 11

The eleventh paper, written by J. Vagners of the University of Washington (Stanford University), presents results on the long-term behavior of close lunar orbiters. A Hamiltonian is presented with the short and medium terms removed. The long-period motion is thus analyzed and results presented.

Paper No. 12

The last paper in this report was written by Hans B. Schechter of Stanford University. This paper presents an analytic study of the three-dimensional stability of motion of a particle near L_4 in a non-linear Earth-moon force field. A linear solar gravitational field distribution is superimposed on the Earth-moon field. The long-period features of the motion of the particle are studied.

INTRODUCTION

Two internal publications authored or co-authored by members of the sponsoring laboratory and in the subject technical fields have appeared since the last compilation. These are listed below with their abstracts.

Miner, W. E., and J. F. Andrus: Necessary Conditions for Optimal Lunar Trajectories with Discontinuous State Variables and Intermediate Point Constraints. NASA TM X-1353, April 1967.

ABSTRACT

The guidance regime for an optimal multi-stage lunar trajectory is derived by applying the mathematics of the calculus of variations as established by Denbow for the generalized problem of Bolza. The steering angle programs for four constant thrust level phases and the times to initiate and terminate the two coast phases are determined in order to place maximum payload into a specified lunar orbit. The problem considered here is to determine an optimal trajectory consisting of six sub-arcs utilizing three vehicle stages on which maximum payload is transported from an exo-atmospheric point near the Earth to a specified lunar orbit. The intermediate point constraints include two points at which stages are separated and mass discontinuities occur, an Earth parking orbit of specified energy and angular momentum magnitude, and four thrust magnitude levels. The Euler-Lagrange equations determine the optimal steering for the thrusting phases and the Denbow transversality equations are used to calculate the discontinuities at the ends of the sub-arcs. This method is applied here and the equations necessary to solve this problem using a high-speed computer are derived.

Hoelker, R. F.: Numerical Studies of Transitions between the Restricted Problem of Three Bodies and the Problem of Two Fixed Centers and the Kepler Problem. NASA TM X-1465, November, 1967.

ABSTRACT

For the comparison of the trajectories of the two fixed-center problem with those of the restricted problem of three bodies, fields of trajectories are numerically computed for six initial position conditions, all starting on the line of masses and perpendicularly to this line. The transition is studied by varying stepwise, for these fields, the rate of revolution of the primaries from zero to the equilibrium rate of circular motion.

For the investigation of the behavior of trajectories in the transition from the Kepler problem to the restricted problem, the Kepler orbits are considered in a coordinate system rotating at the rate determined by the continuation of this system into that of the restricted problem. Transition characteristics are shown on samples of periodic orbits and of trajectory fields. In particular, the continuous transition of a family of Kepler orbits into periodic orbits about the smaller primary as well as about the L_3-libration point is demonstrated.

AN ANALYSIS OF GUIDANCE MODES

By C. G. Pfeiffer
Head, Mathematical Physics Section
Guidance and Analysis Department, TRW Systems,
Redondo Beach, California

NASA Contract NAS 12-593

An Analysis of Guidance Modes[1]

C. G. Pfeiffer[2]

1. Introduction

Guidance theory is concerned with the solution of a two-point boundar, value problem, which arises when one attempts to control[3] the translation motion of a space vehicle so as to attain desired end conditions at mission completion. That is, given the initial conditions, such as the inertial position and velocity of the launch site, and the desired end conditions, such as the orbital elements of the final satellite orbit about the moon or planet, and a description of the characteristics of the rocket engines and/or aerodynamic lift/drag maneuvers which supply the controlling accelerations during the mission, the guidance analyst must design an algorithm for calculating the translational accelerations to be applied as the vehicle moves toward the final time. Then:

Definition 1: Given a mathematical model of the motion of a space vehicle, a description of the translation acceleration which can be commanded by the guidance system, and an estimate of the state of the overall dynamic system, guidance is the task of calculating and executing a realizable acceleration profile which will cause the trajectory of the space vehicle to attain desired end conditions, where

Definition 2: The state of the space vehicle system consists of the position and velocity of the vehicle, the parameters determining the vehicle performance capability, and the parameters determining the gravitational and atmospheric accelerations.

The estimate of the state is obtained from the navigation system, where

Definition 3: Navigation is the task of estimating the state of the space vehicle system from sensed data, such as the first and second integrals of on-board accelerometer data, and/or earth-based tracking data, and/or on-board observation of a celestial reference.

Definition #1 states that guidance encompasses guidance theory as well as the mechanization of the theory. Guidance mechanization usually concerns the guidance theoretician only to the extent that it affects his analytical treatment of the problem. For example, he usually assumes that the attitude control problem can be ignored, where

[1]Acknowledgement: This paper was written under NASA contract NAS 12-593, administered by Electronics Research Center, Cambridge, Massachusetts. The classifications of guidance modes and measures of performance were suggested by Mr. W. E. Miner and Mr. D. H. Schmieder of the Guidance Theory and Trajectory Analysis Branch of ERC.

[2]Head, Mathematical Physics Section, Guidance and Analysis Department, TRW Systems, Redondo Beach, California.

[3]Guidance theory is a special case of final value control theory.

<u>Definition 4</u>: Attitude Control is the task of attaining and stabilizing the vehicle in the attitude configuration called for by the guidance system.

This approach is reasonable for most applications, because the guidance and attitude control response times are usually so different that there is negligible interaction.* On the other hand, the analytical treatment will certainly be dependent upon the functional form of the guidance accelerations, which might be applied in the form of impulsive changes of velocity, realized by thrusting with a relatively high acceleration level for a relatively short time; by starting, throttling, steering and/or shutting off a rocket engine which thrusts for a relatively long period of time, and/or by applying lift and/or drag accelerations during motion in the atmosphere, realized by commanding motion of aerodynamic surfaces. With these considerations in mind, the analyst seeks to design a guidance "mode", where

<u>Definition 5</u>: A guidance mode is a policy for calculating the parameters and functions which will accomplish the guidance task.

Since navigation information will be gathered during the mission in order to update the estimate of the state of the system, a guidance mode must be capable of acting as a real-time feedback final-value control law.

In general, there exists an infinite variety of guidance modes which will accomplish mission objectives. Thus one seeks an optimal guidance policy which satisfies the end conditions while minimizing some performance index, such as engine propellant expenditure, or else one pre-specifies a functional form which is near-optimal. Present practice is to simplify the overall optimization problem by treating it as a sequence of two-point boundary value problems. That is, the overall mission is thought of as a sequence of "phases", usually characterized by the means available for applying the guidance accelerations. The objective of the guidance system for any given phase is to attain an intermediate set of pre-specified end conditions. For example, a guidance phase might consist of transfer by means of relatively high rocket thrust acceleration from a near-earth circular parking orbit to a specified earth-escape hyperbola. (A more detailed description of guidance phases is given in Appendices A and B). Such intermediate end conditions must be obtained by a "targeting" method (discussed in Appendix C). The imposition of these constraints on the overall trajectory leads to a sub-optimal overall guidance law, but, since each phase can be treated individually, the design of appropriate guidance modes is much simplified. In practice, guidance modes for the individual phases are usually quite different in form. Considering also the diverse forms of guidance mechanization employed for the various phases, it is true

*Note that stability is not an important consideration in the guidance problem, for the duration of guidance is finite and short compared to the response time of the dynamic system defined by the translational equations of motion. This is not true for the attitude control problem, indeed, stability is usually the primary design goal.

that essentially different guidance systems are presently used on a given mission.

It seems clear that it would be desirable to design a unified guidance system for future applications, where

Definition 6: A unified guidance system is one capable of guiding all phases of a given mission.

Toward that end, it is the objective of this paper to provide a cursory survey of the present state of guidance theory (see also References 1-4), to describe and classify existing guidance modes, and to define some qualitative measures of their relative performance. The purpose of such a study is to aid in the rational selection and/or development of guidance modes for future missions, to facilitate the synthesis of the chosen modes into a unified guidance concept, and to point out problem areas where further development is required.

2. The Separation of Guidance and Navigation

Implicit in the definitions of guidance and navigation is the assumption that these two problems can be treated separately. That is, in present practice the guidance theorist designs a guidance mode by assuming that the state is known perfectly, and for real-time applications employs the estimated state in place of its true value. This assumption, which is essential to a meaningful discussion of existing guidance modes, requires further clarification.

Strictly speaking, the deterministic derivation of a guidance mode is not correct, for the predicted end conditions which determine the guidance functions become random variables if there are random estimation errors and random systematic disturbances to the trajectory. In effect, the state of the system can no longer be defined simply by position, velocity, and system parameter vectors. Instead, the state must be thought of as the expected value of these quantities plus all the statistical moments of their distribution. In other words, the state can only be described by the conditional probability density function of the state, given the navigation data. From the point of view of guidance optimization theory, the random behavior of the dynamic system implies that there no longer exists a field of solutions which are the characteristics of the deterministic Hamilton-Jacobi partial differential equation. Thus, conceptually at least, the notion of a predictable reference trajectory has to be abandoned.

Simple examples of stochastic control problems (see Appendix D) would seem to indicate that the deterministic guidance analysis is not at all valid for realistic problems. As a practical matter, however, stochastic guidance analysis is not required for those applications where (1) an apriori reference trajectory is available, and (2) the random navigation errors and random systematic errors are small. That is, the deterministic analysis applies when the first variation (or perhaps the first and second variations) about some reference trajectory is the dominant consideration. In a first variation analysis the random errors enter linearly and the deterministic approach can be theoretically justified (see work by Joseph, Lou, and Gunkel). Consideration of the second variation (or second and third variations) does not change this conclusion. It then follows that results similar to the deterministic case are obtained but with additional terms which can be interpreted as non-linear corrections or biases which can be calculated. Such assumptions can be thought of as devices to simplify the difficult computational problem of computing expectations.

The assumptions required to justify deterministic guidance analysis usually apply to those phases of a space mission where continuous guidance accelerations are applied, but stochastic considerations become important when the guidance is applied in the form of a sequence of small velocity impulses at unspecified times. The difficulty in this case is finding the times of application. Present practice for the Ranger-Mariner-Surveyor type of mission (Reference 5) is to pre-specify these times by heuristic or empirical rules, and to calculate the real time corrections with a linearized deterministic rule. The non-linear effects of the corrections are treated by an iterative technique. Stochastic considerations are introduced by employing a non-linear maximum likelihood estimator in the navigation equations, and appropriately modifying the targeted end conditions so as to take into account the statistics of the estimation and systematic errors. Small real time (adaptive) variations in the time of application of the corrections are sometimes allowed. This approach demonstrably works well for many applications.

One concludes, then, that the separate treatment of the guidance and navigation problems can indeed be justified for most present day space missions, and that deterministic guidance analysis is valid, if appropriate approximations are made and if appropriate constraints are placed upon the guidance policy.

3. Description and Classification of Typical Guidance Modes

The primary purpose of this paper is to classify typical guidance modes, and to define qualitative measures of their performance (see Introduction). In this part, various guidance modes will be classified according to the mathematical approximations and/or assumptions introduced in their derivation (Table 1). Although many of the methods discussed have never been used in real time applications, they must be considered as possible modes for future missions.

Class 1: Precise Model of Dynamic System - The guidance mode is based upon a mathematical model representing all known accelerations on the vehicle which are numerically significant.

Class 1.1: Expansion of Solutions - The Class 1 guidance mode generates the control as an explicit function of the state for all states in some region of applicability.

Class 1.1.1: Linear Expansion - The control is a linear function of the state. Examples are:

a) delta guidance - guide to null deviations (δ) from a standard trajectory
b) lamda matrix guidance (Reference 6)
c) second variation guidance (References 7,8)
d) impluse velocity-to-be-gained (References 3,5)
e) steering to velocity-to-be-gained (Reference 3)

Class 1.1.2: Non-Linear Expansion - The control is a non-linear function of the state. Examples are:

a) dynamic programming (Reference 9)
b) solution of Hamilton-Jacobi partial differential equation
c) series expansion of equations of motion
d) series expansion of solutions (Reference 10)
e) series expansion of Hamiltonian (Reference 11)

Class 1.2: Successive Approximations - The class 1 guidance mode generates the control function by successively applying in real time an algorithm which converges after some number of iterations to the optimal control.

Class 1.2.1: Direct Methods - A suitable approximate ordinary extremum problem is defined in which only a finite number (N) of parameters is to be determined, and then the class 1.2 guidance mode generates the control by passing to the limit as $N \to \infty$ in the solution of the approximate problem. Essentially, a minimizing sequence of control functions is constructed, and the desired solution is obtained by a limiting process based upon this sequence (Reference 12, pp 174-175). Some approaches which might be used are:

a) steepest descent (References 13,14)
b) Rayleigh-Ritz method (Reference 12, pp 175-176)
c) method of finite differences (Reference 12, pp 176-177)

Class 1.2.2: Indirect Methods - The class 1.2 guidance mode successively approximates in real time the control which satisfies optimality conditions and the desired end conditions. Examples are:

a) optimal impulsive velocity-to-be-gained (Reference 5)
b) second variation control (References 7,8)
c) sweep method (Reference 15)
d) quasilinearization (Reference 16)
e) quasi-second order approximations (Reference 17)

Class 2: Approximate Model of System Dynamics - The guidance mode is based upon an approximate mathematical model of the vehicle dynamics and/or the gravitational and atmospheric acceleration. Essentially, approximations are introduced into the physical model in order to simplify the mathematical solution of the equations of motion.

Class 2.1: Closed Form - The Class 2 guidance mode yields a closed form solution for the control in terms of the given initial conditions and the desired end conditions.

Class 2.1.1: Approximation of Environmental Accelerations - The derivation of the class 2.1 guidance mode follows from approximating the first and second integrals of the gravitational, drag, and/or lift accelerations between the initial and final times as relatively simple functions of initial and end conditions. Examples are:

a) iterative guidance mode (Reference 18)
b) MIT explicit guidance (Reference 19)

c) TRW explicit guidance (Reference 20)
d) Lewis Research Center explicit guidance (Reference 21)
e) Aerospace explicit guidance (Reference 22)
f) Robbins explicit guidance (Reference 23)

Class 2.1.2: Conic Model - The derivation to the class 2.1 guidance mode is based upon the assumption that the orbit desired at thrust termination is a conic, and that thrust acts for a relatively short duration of time. Examples are:

a) impulsive velocity-to-be gained (Reference 3)
b) steering to velocity-to-be gained (Reference 3)

Class 2.2: Successive Approximation of Closed Form Guidance - The Class 2 guidance mode employs algorithms similar to those of Class 2.1, but the errors due to these approximations are iteratively reduced in real time.

Class 2.2.1: Successive Approximation of Environmental Accelerations - The approximated values of the acceleration integrals are successively improved in real time by evaluating the integrals on the current real-time predicted trajectory. No examples of this technique are known to the author.

Class 2.2.2: Successive Approximation of Conic Model - The effects of neglected terms in the conic model are calculated in real time as time varying perturbations acting on the current approximate trajectory, and these effects are introduced into the guidance equations. This approach is similar to general perturbation techniques well-known in celestial mechanics, but no applications to the real-time guidance problem are known to the author.

4. Measures of Performance

The class of closed form guidance modes has recently received much attention, because one obtains a versatile control law which can treat large perturbations and yet requires relatively little preflight calculation. This approach has limited application, however, because the simplifying approximations required for the derivation cannot be justified in general. For example, at any iteration of the guidance calculation it may be assumed that the gravitational acceleration for the remainder of the flight can be approximated by a constant vector. Such an approximation obviously works well when applied to short powered flight arcs, but for long arcs it can lead to serious degradation of performance. Some one of the Class 1 modes could be used to eliminate such difficulties, but then extensive preflight calculation and storage might be necessary, or,if a linear guidance law is employed, the performance could be poor in the presence of large perturbations. Thus the choice of a guidance mode is often ad hoc, and one must consider many factors. Some measures of performance are:

1. optimality - given that there is a performance index to be minimized, say propellant expenditure, how does the obtained value of the performance index compare to the theoretical minimum?

Table 1: Classification of Possible Guidance Modes

Precise Model of System Dynamics				Approximate Model of System Dynamics			
Expansion of Solutions		Successive approximation of optimal guidance		Closed form		Successive approximation of closed form of guidance	
linear	nonlinear	direct	indirect	Approximation of acceleration integrals	conic model	Approximation of acceleration integrals	conic model
1. delta guidance 2. lamda matrix 3. second variation 4. linearized impulsive guidance 5. impulsive velocity to be gained 6. steering to velocity to be gained	1. dynamic programming 2. solution of Hamilton-Jacobi PDE 3. series expansion of equations of motion 4. series expansion of solutions 5. series expansion of Hamiltonian	1. Rayleigh-Ritz method 2. method of finite differences 3. steepest descent	1. linearized impulsive guidance 2. second variation 3. sweep method 4. quasilinearization 5. quasi-second order approximation	1. iterative guidance mode 2. M.I.T. explicit 3. TRW explicit 4. Lewis explicit 5. Aerospace explicit 6. Robbins explicit	1. impulsive velocity to be gained 2. steering to velocity to be gained	same as closed form, but with real-time revisions of approximations	

2. accuracy - given that approximations are introduced into the derivation and mechanization of the guidance equations, what are the resulting errors in the desired terminal conditions? These errors can be classified according to:

   2.1 approximation errors - due to analytic approximations introduced into the derivation of the guidance equations.

   2.2 computer errors - due to the inaccuracies of the numerical algorithms used to implement the guidance equations (truncation and roundoff).

   2.3 mechanization errors - due to the inability of the vehicle to physically respond to the guidance commands.

   There are also navigation errors, but, insofar as the guidance and navigation problems are separable (i.e., assuming superposition of effects), these errors need not be considered in the design of a guidance mode.

3. region of applicability - what is the range of perturbations which can be adequately treated by the guidance mode?

4. computer factors - what are the real time on-board and/or earth-based computer requirements, in particular, how much storage space is required, what is the length of the computing cycle for each iteration of the guidance equations, and how complex must the computer be?

5. preflight preparation - what is the cost in time and money of preflight preparation of the guidance equations, in particular, how long does it take to prepare the guidance system to accomplish a given mission? (the "quick reaction" problem).

6. flexibility - what are the types of missions which the guidance mode can perform, and how well can it adapt to changes in the mission, such as variations of launch azimuth? (another aspect of the "quick reaction" problem).

7. growth potential - what is potential applicability of the guidance mode to future missions?

5. Conclusions

It is hoped that the classification method and defintiions of measures of performance developed here will be of use in the synthesis of modes (or mode) for a unified guidance system for advanced space missions. Such a system should encorporate the best features of the various modes, and improve upon their limitations. Although existing technology has been adequate for present day missions, some challenging unsolved problems remain in the area of optimal stochastic guidance, especially for the case of impulsively applied guidance corrections (see Part 2). Some interesting results have been obtained (References 24 - 28) but more research is needed.

Appendix A: Description of Guidance Phases

Present practice in guidance system design is to segment the overall trajectory into a sequence of thrusting and coast periods which can be thought of as "phases", where

Definition 7: A guidance phase is a segment of a trajectory, usually characterized by the means available for applying the guidance accelerations, having a distinct guidance objective, i.e., specified end conditions.

The guidance mode in each phase attempts to null errors resulting from the previous phase, plus errors due to any current disturbances, by attaining the end conditions specified for the phase. Thus a guidance system might be called upon to solve in real time many different types of two-point boundary value problems for a wide range of initial conditions. Possible types of guidance phases for advanced missions are:

- High Thrust Continuous Guidance

  Launch vehicle guidance

  a. Initial ascent to altitude
  b. Booster stage
  c. Ascent to orbit
  d. Transfer from parking orbit

  Terminal guidance

  a. Retro into lunar or planetary orbit
  b. Injection into earth satellite orbit
  c. Descent from orbit
  d. Soft landing and hovering

- Low Thrust Continuous Guidance

  Spiral escape from earth

  Earth to target transfer

  a. Lunar
  b. Interplanetary

  Spiral capture by target body

  Continuous orbit adjustment

  a. Earth satellite
  b. Lunar satellite
  c. Planetary satellite
  d. Earth-target transfer orbit

- Impulsive Guidance

  a. Midcourse

  b. Approach

  c. Terminal

  d. Satellite orbit trim

  e. Descent from orbit

  f. Soft landing retro

- Aerodynamic Guidance

  a. Control of lifting reentry

  b. Control of ballistic reentry

  c. Drag brake control

  d. Parachute control

Since there are many types of guidance phases, it is usually the case that more than one guidance mode will be employed during a mission, and more than one command mechanization subsystem will be used. Indeed, the guidance techniques for the various phases are so different that essentially different guidance systems are used during a mission.

Appendix B: Guidance Phases for a Typical Lunar Mission

Ascent Phase

The ascent phase begins at launch and extends to injection into a nearearth circular parking orbit, which might have a standard altitude of 100 n mi above the earth's surface. A typical ascent phase might last 8 to 10 minutes. The objective of the guidance system is to attain circularity (eccentricity equal to zero) at an altitude close to the standard value. Guidance corrections are applied by steering the vehicle with the gimbaled rocket nozzle and by making small changes in the thrust termination time of each rocket stage. The disturbances to the flight path consist of imperfectly applied thrust acceleration and external forces, such as wind and air density variations. The position and velocity of the vehicle are measured by integrating the outputs of accelerometers mounted on an inertially fixed platform within the vehicle, or from ground-based tracking radars, or from both these sources.

Parking Orbit Phase

The parking orbit phase begins at parking orbit injection and extends to the restart of the launch vehicle for the injection phase. Typical parking orbit durations are 1 to 20 minutes (4 to 80° of coast arc) but they can be indefinitely long. There are usually no guidance corrections required during this phase, but some vernier adjustment of the errors remaining from the ascent phase might be made. The disturbances to the flight path are negligibly small for short coast arcs, but otherwise atmospheric drag becomes important. The position and velocity of the vehicle are determined as in the ascent phase, but celestial observations can be incorporated if the parking orbit is sufficiently long. A rendezvous and docking of two or more vehicles may occur in this phase, the purpose being to assemble a spacecraft capable of completing the remainder of the mission. The rendezvous does not alter the essential character of the guidance problem and will not be discussed here.

Injection Phase

The injection phase begins at restart of the launch vehicle and extends to injection into earth-moon transfer orbit (which is an ellipse relative to the earth with eccentricity of 0.987 for a 66 hour transfer). The duration of the injection phase is typically 2 to 3 minutes. The objective of the guidance system is to attain a transfer orbit which will cause the spacecraft to impact the desired target point on the moon. The corrections are made as in the ascent phase. The disturbances to the flight path are primarily due to imperfectly applied thrust acceleration. The position and velocity of the vehicle are determined as in the ascent phase.

Midcourse Phase

The midcourse phase begins at injection and extends until the spacecraft enters the "sphere of influence" of the moon, a point which is not precisely defined but is approximately 60,000 km from the moon. The duration of a typical midcourse phase is roughly 50 hours. The spacecraft is separated from the launch vehicle during this period. The primary objective of the guidance system is to correct for errors in the injection phase, thus providing a vernier adjustment. There are in addition some small disturbances to the flight path to consider, such as solar winds, leaking gas jets in the attitude control system, and errors in the assumed values of the physical constants which define the mathematical model used to construct the standard trajectory. The orbit is determined from celestial sightings and/or earth-based radar data. The guidance corrections are performed by applying short-duration impulses of acceleration (on the order of a minute long) with a small rocket engine so as to achieve "delta functions" of velocity. The magnitude of the correction is determined by the duration of the thrusting, and the desired direction is attained by properly pointing the spacecraft. One or more corrections might be made, the first no sooner than 5 hours after injection so as to allow time to determine the orbit, and others (usually not more than two) as required to null errors in the previous correction. The total impulse applied in the midcourse phase depends primarily on the injection error, but is typically less than 100 m/sec.

Approach Phase

The approach phase begins when the spacecraft enters the sphere of influence of the moon and extends until just prior to beginning of the terminal phase, a period of typically 15 hours. The objective of the guidance system, the disturbances to the flight path, and the techniques for determining the orbit and applying the guidance corrections are the same as in the midcourse phase. The trajectory is a moon-centered hyperbola with a hyperbolic excess velocity of typically 1.0 to 1.2 km/sec. Two or more corrections will probably be made, based on orbit determination measurements which sense the position and/or velocity of the spacecraft relative to the moon. Examples of such observations would be on-board sightings of the angles between the target center and certain stars and/or measurement of the change in spacecraft speed as it is acted upon by the moon's gravitational attraction. It is the gathering of this target-relative type of orbit information which distinguishes the approach phase. Since the ultimate mission accuracy very likely will depend on this information, the approach guidance phase is one of the most important of all. It supplies the final vernier corrections to the orbit.

Terminal Phase

The terminal phase begins at the completion of the last approach correction and extends through the final thrusting required to complete the mission, which might be a retro-braking into satellite orbit, a direct descent to the lunar surface, or a combination of these two maneuvers in order to descend to the surface from parking orbit. Integrated accelerometer data would be used during the thrusting periods, the initial conditions being obtained from the orbit parameters estimated during the approach phase. Celestial measurements and/or earth-based tracking data would be employed, if possible, during the coast periods. Only small impulsive corrections made would be during the parking orbit, if there is one. Thus the terminal phase is similar to the ascent-to-injection phases, with appropriately modified guidance objectives.

Appendix C: The Targeting Problem

Although the objectives of the guidance phases are different, it is obviously necessary that they be compatible and lead to the satisfaction of ultimate mission objectives. Thus an important aspect of guidance analysis is the targeting problem, where:

Definition 8: Targeting a given guidance phase is the task of analytically and/or numerically specifying the objectives of that phase.

Thus targeting is concerned with the practical task of piecing together the solutions of sequence of two-point boundary value problems so as to devise an overall solution of the complete problem. The targeting problem is almost synonomous with the guidance theory problem to analysts primarily concerned with guidance maneuvers which take the form of velocity impulses, while analysts concerned with continuous thrusting think of targeting in terms of specifying end conditions. In the terminology of optimization theory, targeting may be thought of as the task of specifying the transversality conditions for any guidance phase, given that the trajectory has been segmented into phases.

The conic formulae are used extensively in targeting, for motion during a coast period in a drag-free environment can usually be closely approximated, with perhaps some empirical correction terms, by the solutions of "patched" two-body problems. Thus for guidance purposes a closed form solution is valid in these segments of the trajectory, and the objectives of a given guidance phase can be stated as attaining a certain combination of orbital elements. A dynamic programming argument can be used to develop the sequence of desired elements for all phases by working backward from mission termination. For example, the objective of a retro thrust maneuver to obtain injection into a terminal satellite orbit about a planet can be specified in terms of the elements of that orbit. The objectives of the approach phase can be specified in terms of the elements of an approach hyperbola which will be optimal for the retro phase. The objectives of the midcourse phase can be specified in terms of the elements of a heliocentric transfer ellipse which will yield an optimal approach hyperbola. The objectives of the near-earth injection phase can be specified in terms of the elements of the earth-escape hyperbola which will yield an optimal heliocentric transfer ellipse. Lastly, the objectives of the initial ascent from the launch pad can be specified in terms of the elements of the near-earth parking orbit which yields an optimal injection phase.

The notion of a "patched" conic is not a precise one, for the actual trajectory is continuously attracted by many bodies. The guidance analyst does not consider the conics to be joined at fixed points on the trajectory, however, but instead they are "asympototically matched" in order to yield a much better approximation of the true motion. That is, the target planet can be considered massless for the purpose of injection and midcourse guidance analysis, and the position and velocity at closest approach to the target can then be used to determine the asymptote of the approach hyperbola (reference 29). The magnitude of the position vector at closest approach is the impact parameter (b), and the velocity vector is the hyperbolic excess velocity ($\bar{v}_\infty$). The energy and angular momentum of the approach hyperbola are then given by $c_3 = |\bar{v}_\infty|^2$ and $c_1 = b|\bar{v}_\infty|$, respectively. The errors introduced by such an approximation are due to the differences in gravitational acceleration of the target and spacecraft caused by the attraction of non-target masses acting during the approach phase (such as the sun), and are usually negligibly small compared to other sources of guidance system error.

It is usually necessary to predict and/or control the time of flight to closest approach to the target. Consistent with the notion of a massless target, one might take the x_1 coordinate axis in the direction of the target-spacecraft relative velocity ($\bar{v}_\infty$) as determined on the standard trajectory at the standard time of closest approach (t_{fs}). The predicted first-order change in impact time then becomes

$$\delta t_f = \frac{\delta x_1(t_f)}{|\bar{v}_\infty|}$$

This approach to the problem is well-suited to optimization analysis, for the minimum time trajectory is obtained by minimizing x_1 at the <u>fixed</u> time t_f. Another more commonly used technique is to employ the so-called linearized-time-of-flight, which is the time of closest approach to a massy target corrected for the non-linear effects of the impact parameter, given by (references 30 and 31)

$$t_L = t(\text{closest approach}) + \frac{b}{|\bar{v}_\infty|} \ln e$$

where e is the eccentricity of the approach hyperbola. It can be demonstrated analytically and numerically that t_L behaves almost as a linear function of the midcourse correction components.

Appendix D: A Simple Stochastic Control Problem

The discussion of guidance modes presented here deals with the deterministic guidance problem, assuming that random navigation and systematic errors can be treated separately. As indicated in Part 2, this is a reasonable assumption for many applications, but does not hold for problems where the random errors are large, or where the guidance corrections are applied on the form of impulses at times which are not a priori specified. In these situations the deterministic approach is not appropriate, and a stochastic guidance law should be devised.

The stochastic control problem can be illustrated by a simple example. Suppose that at the initial time t_o the final state $x_1(t_f)$ is known to be of the form

$$x_1(t_f) = g\ (u,\ \alpha)$$

where u is some scalar control parameter, and α is some scalar parameter characterizing the motion between t_o and t_f. For example, α might be the initial condition $x_1(t_o)$, or the magnitude of a perturbing acceleration acting between t_o and t_f. Suppose the navigation system has provided an estimate of α, denoted by α^*, but there is an error in this estimate, denoted by $\varepsilon=(\alpha^*-\alpha)$. Suppose this unknown error is a zero mean, Gaussian random variable with variance σ^2 over the ensemble of all similar experiments. The problem is to choose the u which in some sense minimizes x_1. If the estimates were perfect ($\varepsilon=o$), one would seek a u^o such that

$$\frac{\partial g}{\partial u}\ (u^o,\ \alpha^*) = 0$$

In the stochastic case, however, the error in the estimate can be arbitrarily large, so one must deal with the statistical expectation (E [$\cdot$]) of the derivative. This might be expressed in the form of a Taylor series as

$$0 = E\left[\frac{\partial g}{\partial u}\ (u^o,\alpha)\right] = E\left[\left(\frac{\partial g}{\partial u}\right) + \left(\frac{\partial^2 g}{\partial u\ \partial\alpha}\right)\varepsilon + \frac{1}{2}\left(\frac{\partial^3 g}{\partial u\ \partial\alpha^2}\right)\varepsilon^2 \right.$$

$$+ \frac{1}{3!}\left(\frac{\partial^4 g}{\partial u\ \partial\alpha^3}\right)\varepsilon^3 + \frac{1}{4!}\left(\frac{\partial^5 g}{\partial u\ \partial\alpha^4}\right)\varepsilon^4$$

$$\left. + \text{ higher order terms in } \varepsilon^n\right]$$

$$= \left(\frac{\partial g}{\partial u}\right) + \frac{1}{2}\left(\frac{\partial^3 g}{\partial u\ \partial\alpha^2}\right)\sigma^2 + \frac{3}{4!}\left(\frac{\partial^5 g}{\partial u\ \partial\alpha^4}\right)\sigma^4$$

$$+ \text{ [higher order terms in } \sigma^2]$$

where the coefficients of the Taylor series are evaluated as functions of u^o and α^*, and properties of a Gaussian distribution have been used in computing the expectation (i.e., the expected value of the odd moments are zero, and the expected value of the even moments are expressible in terms of σ^2). Thus it can be seen that the statistics of the errors in the estimate become inseparably mixed into the optimization problem, and it could be difficult to find u^o. The Taylor series method might not be the best approach, expecially for states of higher dimension. In any case, one is conceptually faced with solving an infinite number of optimization problems corresponding to every possible value of ε, and choosing a weighted (according to the probability of occurrence of the values of ε) average of the solutions.

A still more subtle problem arises if it is desired to choose two guidance parameters (u_1, u_2) to minimize

$$x_1(t_f) = g_1(u_1, u_2, \alpha)$$

subject to

$$x_2(t_f) = g_2(u_1, u_2, \alpha) = \text{given value}$$

Analogous to the deterministic case and the previous example, one is tempted to minimize

$$E\ [g_1(u_1, u_2, \alpha) + \nu\ g_2\ (u_1, u_2, \alpha)]$$

where ν is a Lagrange multiplier. Conceptually this corresponds to solving an infinite number of optimization problems where the end condition is satisfied each time, and the Lagrange multiplier has to be treated as a random variable. There is no reason to expect that suc solutions exist, however. Alternatively, one seeks to minimize

$$E\ [g_1(u_1, u_2, \alpha)] + \nu\ E[g_1(u_1, u_2, \alpha)]$$

where ν is a fixed constant. In this case the end conditions are satisfied "on the average", that is, the expected value of the end conditions satisfy the constraint but individual members of the ensemble generally do not.

It appears that stochastic control considerations analogous to those discussed in this simple example will arise in the case of impulsively applied guidance corrections. Some qualitative and approximate solutions of such problems have been obtained (references 24 - 28), but more work in this area is clearly indicated.

References

1. C. G. Pfeiffer, "Guidance of Unmanned Lunar and Interplanetary Spacecraft", Celestial Mechanics and Astrodynamics, edited by Victor G. Szebehely, (Academic Press, New York, 1964), pp. 259-279.

2. C. G. Pfeiffer, "Guidance Analysis," Lunar Missions and Exploration (John Wiley and Sons, Inc., New York, 1964), pp. 276-307.

3. R. H. Battin, Astronautical Guidance, (McGraw-Hill, New York, 1964).

4. R. F. Hoelker and W. E. Miner, "Introduction into the Concept of the Adaptive Guidance Mode," Aeroballistics Internal Note No. 21-60, Marshall Space Flight Center (December 28, 1960).

5. A. R. M. Noton, E. Cutting, and F. Barnes, "Analysis of Radio-Command Midcourse Guidance," Technical Memorandum 32-38, Jet Propulsion Laboratory, Pasadena (June 1959).

6. A. E. Bryson and W. F. Denham, "Multivariable Terminal Control for Minimum Mean Square Deviation from a Nominal Path," Proceedings of the IAS Symposium on Vehicle Systems Optimization, (Garden City, N. Y., November 1961).

7. J. V. Breakwell, J. L Speyer, and A. O. Bryson, "Optimization and Control of Nonlinear Systems Using the Second Variation," SIAM J. Control 1, (1963) pp 193-223.

8. H. J. Kelley, "Guidance Theory and Extremal Fields, "Trans. Inst. Radio Engineers 7, (1962) pp 75-82.

9. S. E. Dreyfus, Dynamic Programming and the Calculus of Variations, RAND Report R-441-PR (August 1965).

10. W. E. Miner, D. H. Schmieder, and N. J. Braud, "Path Adaptive Mode for Guiding Space Flight Vehicles," Progress in Astronautics and Rocketry, Vol. 8, Academic Press, 1962).

11. W. E. Miner, B. D. Tapley, W. F. Powers, "The Hamilton-Jacobi Method Applied to the Low-Thrust Trajectory Problem," paper presented at XVIII IAF Conference, Belgrade, Yugoslavia, (September 29, 1967).

12. R. Courant and D. Hilbert, Methods of Mathematical Physics, (Interscience Publishers, Inc., New York, 1953).

13. H. J. Kelley, Gradient Theory of Optimal Flight Paths, ARS Semi-annual Meeting, Los Angeles, California, May 9-12, 1960 ARS Journal, (October 1960).

14. A. E. Bryson, and W. F. Denham, A Steepest-Ascent Method for Solving Optimum Programming Problems, Report BR-1303, Raytheon Missile and Space Division, (August 10, 1961).

15. S. R. McReynolds, "A Successive Sweep Method for Constructing Optimal Trajectories," PhD. dissertation, Harvard University, Cambridge, Massachusetts (September 1966).

16. R. McGill and P. Kenneth, "Solution of Variational Problems by Means of a Generalized Newton-Raphson Operator", Grumman Aircraft Engineering Report GRD-100A, April 1964.

17. C. G. Pfeiffer, "A Successive Approximation Technique for Constructing a Near-Optimal Guidance Law," paper presented at XVIII IAF Conference, Belgrade, Yugoslavia, (September 29, 1967).

18. I. E. Smith, "General Formulation of the Iterative Guidance Mode," NASA Report TM X-53414, (22 March 1966).

19. G. W. Cherry, "A General, Explicit, Optimizing Guidance Law for Rocket Propelled Spaceflight," Proceedings of the AIAA/ION Astrodynamics Guidance and Control Conference, (24 August 1964).

20. "Centaur Explicit Guidance Equation Study Final Report," TRW Technical Report No. 08768-6002-R000, Prepared for NASA Lewis Research Center, Under Contract No. NAS3-3231, Amendment 19 (17 January 1967).

21. F. Teren, "Explicit Guidance Equations for Multistage Boost Trajectories," NASA Report TN D-3189 (6 August 1965).

22. F. M. Perkins, "Derivation of Linear-Tangent Steering Laws," Air Force Report No. SSD-TR-66-211, Aerospace Report No. TR-1001(9990)-1, (November 1966).

23. H. M. Robbins, "Optimal Steering for Required Velocity Guidance", Navigation, 12, 1965, pp 35.

24. R. J. Orford, "Optimal Stochastic Control Systems," J. Mathematics Anal. Applications 6, (1963) pp 419-429.

25. A. Rosenbloom, "Final Value Systems With Total Effort Constraint", Proceedings of the First International Federation of Automatic Control (Butterworth Scientific Publications, Ltd., London, 1960).

26. J. V. Breakwell, "The Optimum Spacing of Corrective Thrusts in Interplanetary Navigation," Mathematics in Science and Engineering (Academic Press, Inc., New York, 1962), Vol. 5.

27. C. G. Pfeiffer, "A Dynamic Programming Analysis of Multiple Guidance Corrections of a Trajectory," AIAA Journal, Vol. 3, No. 9 (September 1965) pp. 1674-1681.

28. D. W. Curkendall, "Monte Carlo Simulation of an Adaptive Policy for Multiple Impulse Correction of the Trajectory of a Spacecraft," AAS Preprint (February 1965).

29. W. A. Kizner, A Method of Describing Miss Distances for Lunar and Interplanetary Trajectories, External Publication 674, Jet Propulsion Laboratory, Pasadena (August 1959).

30. J. O. Malloy, "Non-linear Behavior of Time of Flight to Closest Approach," Unpublished Jet Propulsion Laboratory Internal Memorandum, Pasadena (1963).

31. J. V. Breakwell and L. M. Perko, "Matched Asymptotic Expansions, Patched Conics, and the Computation of Interplanetary Trajectories," Progress in Astronautics, (Academic Press, Inc., New York, 1966), Vol. 17 pp 159-182.

OPTIMAL GUIDANCE FROM HYPERBOLIC TO CIRCULAR ORBITS

By T. N. Edelbaum
Analytical Mechanics Associates, Inc.
Cambridge, Massachusetts

OPTIMAL GUIDANCE FROM HYPERBOLIC TO CIRCULAR ORBITS

T. N. Edelbaum
Analytical Mechanics Associates, Inc.
Cambridge, Massachusetts

ABSTRACT

An approximate analytic solution is developed for minimum fuel guidance from an arbitrary point on a hyperbolic orbit into a specified circular orbit. The hyperbola must lie close to the plane of the circular orbit and its periapsis radius must be close to the radius of the circular orbit. Optimization of the midcourse impulse, the finite-thrust terminal burn, and of both maneuvers in combination is considered. The particular problem treated is intended as a simple example of a new, unified guidance technique.

Presented at the Colloquium on
Advanced Problems and Methods for Space Flight Optimization
Institute of Mathematics, University of Liege, Belgium
June 19-23, 1967

NOMENCLATURE

c	Exhaust velocity
e	Eccentricity of nominal approach hyperbola
f	True anomaly
F	Thrust
m	Mass
M_1	First moment of acceleration during final burn
M_2	Second moment of acceleration during final burn
R	Radius
t	Time
u	Corrective velocity change
V_R	Relative velocity at nominal target interception
x	Circumferential component of position at nominal target
y	Radial component of position at nominal target
z	Out-of-plane component of position at nominal target
α	Thrust angle with local horizontal
β	Thrust angle out of orbit plane
μ	Gravitational constant of planet
ω	Rate of change of α
Ω	Rate of change of β

Subscripts

c	Critical
cy	Component of critical plane correction in orbit plane
cz	Component of critical plane correction out of orbit plane
nc	Noncritical
p	Periapsis
0	Beginning of terminal burn
1	Centroid of terminal burn
2	End of terminal burn

OPTIMAL GUIDANCE FROM HYPERBOLIC TO CIRCULAR ORBITS

INTRODUCTION

In recent years, there has been increasing interest in the use of optimization theory for the guidance of space vehicles. Much of this work has been deterministic and is based upon the idea of flying a minimum fuel trajectory from the vehicle's current state to the desired final state. The early work in this field either linearized both the state and the control around the nominal trajectory (second variation guidance, Refs. 1, 2) or else numerically fit the control to a pencil of optimal trajectories which achieved the desired final state (path adaptive guidance, Ref. 3). A more recent trend has been to use analytical solutions to optimal trajectories for various approximations to the full equations of motion (Refs. 4-6). The latter technique is now widely used for booster guidance but space-craft guidance is still largely based upon various ad hoc approximations (Refs. 7, 8). What is needed is a unified theory for the optimal guidance of space vehicles which can handle various missions and phases of flight. Such a theory might be developed by the onboard generation of nominal trajectories with the use of neighboring extremals for some segments of the trajectory and analytical solutions of approximations to the full problem for the other segments of the trajectory. While the concept of a neighboring extremal may be used (as in second variation guidance), it will not be permissible to linearize the control (as is done in second variation guidance) because many of the portions of the trajectory have no control.

The present study represents a first step in the development of a unified theory of optimal guidance of space vehicles. The theory is for minimum fuel deterministic guidance of high thrust vehicles such as an advanced kick stage. It treats both midcourse and terminal guidance in a unified fashion for variable-time-of-arrival missions. The primary application would be to orbiter missions and to rendezvous missions.

The theory is developed by considering a particular example which is simple enough to allow explicit development of the equations but still reflects many of the difficulties to be encountered in more complicated examples. The particular example treated is guidance in a central gravitational field from a hyperbolic approach condition into a desired given terminal circular orbit. The time for this operation is open and the initial perturbations from the nominal approach hyperbola are taken to be small. The nominal approach hyperbola is tangent to, and in the plane of the circular orbit. However, the perturbations around this nominal trajectory may occur in three dimensions.

The analysis is divided into three parts. First, a general solution is given for an optimal midcourse correction of a trajectory which has a finite terminal impulsive maneuver. The special case of the problem considered in this paper is treated in this section. The second part of the paper is concerned with the motion during a finite thrust terminal burn. On the basis of this analysis, due largely to Robbins (Ref. 9), a guidance logic is proposed for guidance during the terminal phase. In the third part of the paper, the terminal maneuver and midcourse maneuver are considered together so that an overall optimization of the combined corrections may be carried out.

ANALYSIS

I - Optimum Midcourse Correction

Most existing theories of optimal midcourse guidance are for cases where all of the corrections are infinitesimal corrections in the neighborhood of an unpowered freefall trajectory (e.g. Refs. 10-13). It is necessary to extend this theory to the case where the nominal trajectory may contain powered arcs of finite magnitude. This section will consider the simplest case of such a problem. For this case, the nominal trajectory has a single finite impulse at its end and represents an optimum variable-time-of-arrival rendezvous or orbit transfer. At some time, it is discovered that the actual trajectory has departed by a small amount from the nominal trajectory and a correction must be made in order to insure that the original objectives of the mission will be met. In many cases, only a single midcourse correction will be necessary. This will be true for the particular example treated in this paper and is the only case that will be considered herein. The three components of the midcourse velocity correction will produce changes in three components of the terminal position and in three components of the terminal velocity. A special co-ordinate system has been used for the analysis of variable-time-of-arrival position guidance which is also useful in the more general problem under consideration. This co-ordinate system makes use of a direction known as a non-critical direction (Refs. 10-11). A small velocity impulse in a non-critical direction will produce changes in the terminal position which are parallel to the relative velocity between the vehicle and its target at the nominal arrival time. In the case of an orbiter mission, this relative velocity vector may be taken as the direction of the terminal impulse. At right angles to the non-critical direction is the critical plane in which the position deviations which are orthogonal to the terminal impulse are corrected.

The total velocity correction in addition to the nominal characteristic velocity of the nominal trajectory is given by Eq. 1.

$$\Sigma\delta V = \sqrt{u_c^2 + u_{nc}^2} - \frac{\partial V_R}{\partial u_{nc}} u_{nc} - \frac{\partial V_R}{\partial u_c} u_c \tag{1}$$

This is a first order expression and considers both the velocity change in the midcourse impulse and the corresponding changes in the terminal impulse. The first term is simply the Euclidean norm of the velocity changes in the critical and non-critical directions. The magnitude of the velocity change in the critical plane will be determined by the requirement that the position deviation normal to the direction of the terminal impulse must be reduced to zero. The component of velocity in the non-critical direction will be used to reduce the relative velocity and the magnitude of the terminal velocity impulse. The second term represents the change in the magnitude of the terminal impulse due to a small impulse in the non-critical direction, while the third term represents the reduction in magnitude of the terminal impulse due to a small impulse in the critical direction. As we are only considering first order terms, only the changes in velocity parallel to the finite terminal impulse need be considered. The optimum magnitude of the component of the midcourse impulse in the non-critical direction may be found by differentiating Eq. 1 with respect to this velocity component and setting the derivative equal to zero. Eq. 1 always possesses a single minimum if the nominal trajectory is optimal. The optimum magnitude of the component of velocity in the non-critical direction is given by Eq. 2, while the corresponding minimum cost due to the trajectory correction is given by Eq. 3.

$$u_{nc}^* = \frac{\frac{\partial V_R}{\partial u_{nc}} \left| u_c \right|}{\sqrt{1 - \left[\frac{\partial V_R}{\partial u_{nc}} \right]^2}} \tag{2}$$

$$\Sigma\delta V^* = \sqrt{1 - \left[\frac{\partial V_R}{\partial u_{nc}} \right]^2} \left| u_c \right| - \frac{\partial V_R}{\partial u_c} u_c \tag{3}$$

Eqs. 1 through 3 represent the general linear solution for the optimum magnitude and direction of a single midcourse impulse at a specified location when there will be a large terminal impulse and when the transfer time is open. For the deterministic case considered herein, such an impulse at the earliest possible time will normally provide an optimum correction. For example, this is always true for the transfer from a hyperbola to a nearly coplanar circle considered herein. For any particular case, the optimality of a single correction may be checked by calculating the primer vector (Refs. 14, 15) along the unpowered trajectory. If the primer vector exceeds the magnitude of unity at any point other than the initial and terminal points, an additional impulse will be required for an optimum trajectory. Ref. 15 suggests a method for calculating this multiple impulse correction.

It should be noted that this optimum midcourse correction depends only upon the position deviations normal to the direction of the final impulse. It does not depend upon any of the velocity deviations at the terminal point on the trajectory. The velocity deviations normal to the terminal impulse do not affect the fuel consumption to first order and are neglected, while the deviation parallel to the final impulse is corrected by the final impulse and does contribute linearly to the total characteristic velocity.

For the particular example treated herein, the nominal trajectory is a hyperbola which is coplanar with, and tangent to, the terminal circular orbit. The deviations from the nominal position and velocity of the periapsis will be analyzed in a Cartesian inertial coordinate system whose x-axis is aligned with the direction of the circumferential final impulse. The y-axis will be radial and the z-axis will point out of the orbit plane. The rates of change of the significant components of the periapsis position and velocity with respect to the magnitude of a midcourse impulse are given by Eqs. 4 through 6.

$$\frac{dy_p}{du} = \sqrt{\frac{R_p^3}{\mu(e+1)}} \left[(1-\cos f) \frac{2+e+e\cos f}{1+e\cos f} \cos\alpha - \sin f \sin\alpha \right] \cos\beta \tag{4}$$

$$\frac{dz_p}{du} = \sqrt{\frac{R_p^3(e+1)}{\mu}} \frac{\sin f}{1+e\cos f} \sin\beta \tag{5}$$

$$\frac{d\dot{x}_p}{du} = \left[(1+e\cos f)\cos\alpha + e\sin f\sin\alpha \right] \frac{\cos\beta}{e+1} \tag{6}$$

The noncritical direction is in the plane of the circular orbit and is given by Eq. 7.

$$\tan\alpha_{nc} = \sqrt{\frac{1-\cos f}{1+\cos f}} \frac{2+e+e\cos f}{1+e\cos f} \tag{7}$$

This is the direction of thrust which will produce no change in the periapsis radius to first order. It is now convenient to reference angles in Eq. 6 to the non-critical direction to yield Eq. 8.

$$\frac{d\dot{x}_p}{du} = \frac{\sqrt{1+\cos f}\,(e+1)\cos(\alpha-\alpha_{nc}) - 2\sqrt{1-\cos f}\,\frac{1+e\cos f}{e+1}\sin(\alpha-\alpha_{nc})}{\sqrt{(e+1)^2(1-\cos f)+2(1+e\cos f)(2+e-\cos f)}} \cos\beta \tag{8}$$

From Eq. 8, the partial of the terminal impulse with respect to a midcourse impulse in the non-critical direction can be found as Eq. 9.

$$\frac{\partial V_R}{\partial u_{nc}} = \frac{\partial \dot{x}_p}{\partial u_{nc}} = \frac{\sqrt{1+\cos f}\,(e+1)}{\sqrt{(e+1)^2(1-\cos f)+2(1+e\cos f)(2+e-\cos f)}} \tag{9}$$

Eq. 9 varies monotonically with distance from the periapsis, between the two limits given by Eq. 10.

$$\sqrt{\frac{e-1}{e+1}} \leq \frac{\partial V_R}{\partial u_{nc}} \leq 1 \tag{10}$$

The left-hand limit is the value at an infinite distance while the right-hand limit is the value at periapsis. The magnitude of this quantity is always less than unity, as it must be if the single correction is to be optimum. By using Eqs. 2, 4, and 5, the three components of the optimal midcourse correction may be found as Eqs. 11, 12, and 13.

$$u_{cy} = -y_p \sqrt{\frac{\mu(e+1)}{R_p^3(1-\cos f)}} \; \frac{1+e\cos f}{\sqrt{(e+1)^2(1-\cos f)+2(1+e\cos f)(2+e-\cos f)}} \tag{11}$$

$$u_{cz} = -z \sqrt{\frac{\mu}{R_p^3(e+1)}} \; \frac{1+\cos f}{\sin f} \tag{12}$$

$$u_{nc} = \frac{\sqrt{1+\cos f}\,(e+1)\sqrt{u_{cy}^2+u_{cz}^2}}{\sqrt{2(1+e\cos f)(2+e-\cos f)-2(e+1)^2\cos f}} \tag{13}$$

Finally, the total additional cost of the midcourse impulse and the change in the magnitude of the terminal impulse due to position deviations in the periapsis is given by Eq. 14.

$$\Sigma\delta V = \sqrt{\frac{2(1+e\cos f)(2+e-\cos f)-2(e+1)^2\cos f}{2(1+e\cos f)(2+e-\cos f)+(e+1)^2(1-\cos f)}}\sqrt{u_{cy}^{\,2}+u_{cz}^{\,2}}$$

$$+\frac{2\sqrt{1-\cos f}\;\dfrac{1+e\cos f}{e+1}\;u_{cy}}{\sqrt{2(1+e\cos f)(2+e-\cos f)+(e+1)^2(1-\cos f)}} \tag{14}$$

It is interesting to compare the fuel consumption due to this optimum correction strategy with the fuel consumption of a technique which has been used previously, namely the minimization of the magnitude of only the midcourse impulse. Figure 1 illustrates the total magnitude of velocity corrections due to a midcourse correction impulse at the illustrated locations and in the illustrated directions. The magnitude of each arrow represents the total corrective velocity (Eq. 14) necessary to correct a unit displacement in periapsis altitude. The arrows above the hyperbola are the corrections necessary to raise the periapsis while those below the hyperbola are those necessary to depress the periapsis. At each point there are two arrows which are diametrically opposite. These arrows are in the critical direction which minimizes the magnitude of only the midcourse impulse. The arrows are not of the same length because the change in the magnitude of the terminal impulse due to raising or to lowering periapsis is of opposite sign. The two shorter arrows at each point represent the optimum corrections. The directions of these arrows are symmetric with respect to the non-critical direction which is orthogonal to the critical direction. The figure shows that the optimum correction strategy always saves fuel relative to the other strategy and that the relative fuel saving becomes quite large as the vehicle gets closer and closer to periapsis.

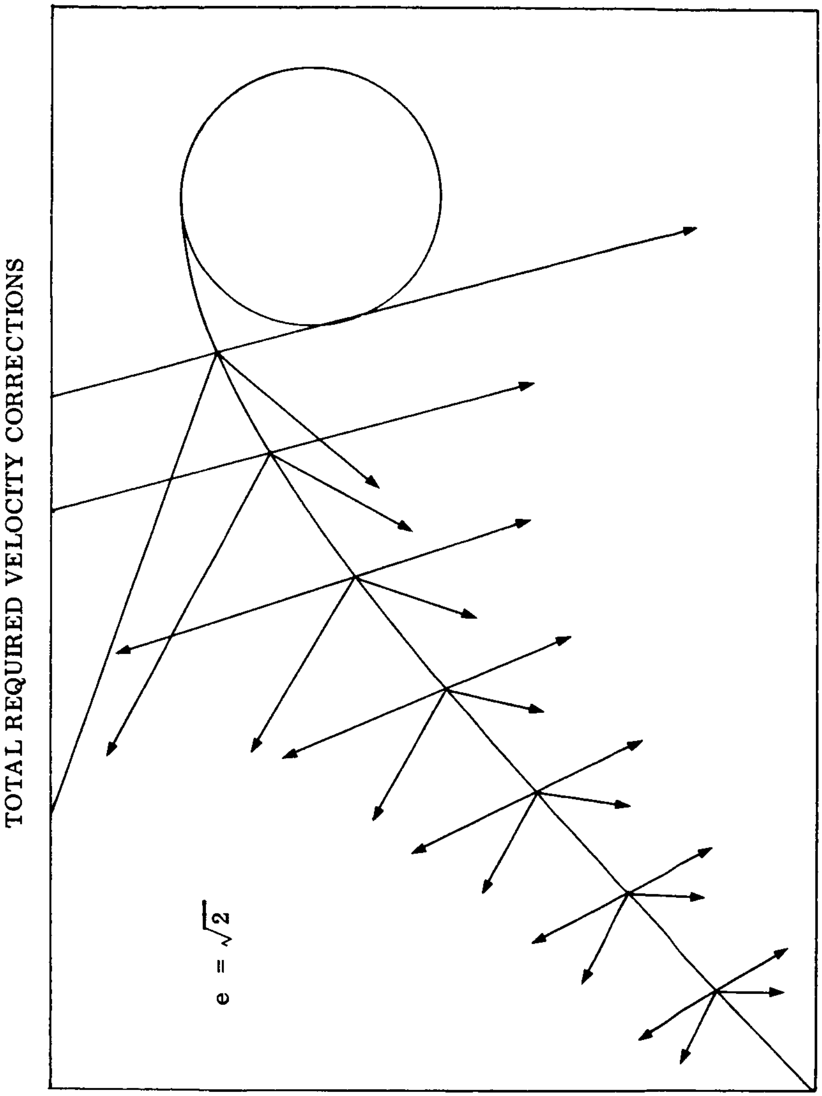


Figure 1

II - Terminal Maneuver

The finite thrust terminal maneuver for injection into the circular orbit cannot be approximated as an impulse because of its relatively long duration. The guidance for this terminal maneuver is developed by utilizing an approximate solution of the equations of motion. This approximate solution is a direct application of a theory due to Robbins (Ref. 9). It represents a second order expansion of the impulsive solution for the optimal trajectory in terms of inverse powers of the average acceleration. Robbins' theory can theoretically treat any thrust acceleration schedule because it is based upon the moments of the acceleration about a centroid time. For the practically important case of constant thrust, the moments about the centroid of the finite thrust burn are given by Eqs. 16-19, while the characteristic velocity of the maneuver is given by the standard rocket equation, 15.

$$\Delta V \equiv \int_{t_0}^{t_2} \frac{F}{m}\, dt = c \,\ell n \frac{m_0}{m_2} \tag{15}$$

$$M_1 \equiv \int_{t_0}^{t_2} \frac{F}{m} (t - t_1)\, dt = 0 \tag{16}$$

$$t_1 - t_0 = \frac{cm_0}{F} \left[1 - \frac{c}{\Delta V} \left(1 - e^{-\frac{\Delta V}{c}} \right) \right] \tag{17}$$

$$t_2 - t_1 = \frac{cm_0}{F} \left[\frac{c}{\Delta V} \left(1 - e^{-\frac{\Delta V}{c}} \right) - e^{-\frac{\Delta V}{c}} \right] \tag{18}$$

$$M_2 \equiv \int_{t_0}^{t_2} \frac{F}{m} (t - t_1)^2 \, dt$$

$$= \frac{(t_2 - t_0)^2 \Delta V}{12} \left[1 - \frac{1}{60} \left(\frac{\Delta V}{c} \right)^2 + \frac{1}{1920} \left(\frac{\Delta V}{c} \right)^4 + \ldots \right] \tag{19}$$

Robbins' analysis is carried out in an inertial Cartesian coordinate system whose origin is at the location of the nominal impulsive burn and whose x-axis is aligned with the direction of the impulse. For the present problem, the impulse is circumferential and that will be the direction of the x-axis. The y-axis will be taken to be radial and the z-axis will be taken to be normal to the first two axes. The analysis in this rectangular co-ordinate system is then developed by ignoring all powers of the inverse of the average acceleration higher than the second. Robbins shows that, for the time open case, this assumption implies that small angle formulas may be used, that the optimal thrust direction is a linear function of time, but that the gravity gradient in the nominal direction of the impulse must be considered in the analysis. With these assumptions, the equations of motion are given by Eqs. 20-22.

$$t_1 \equiv 0$$

$$\ddot{x} + \frac{\mu}{R^3} x = - \frac{F}{m} \left[1 - \frac{(\alpha_1 + \omega t)^2 + (\beta_1 + \Omega t)^2}{2} \right] \tag{20}$$

$$\ddot{y} + \frac{\mu}{R^2} = - \frac{F}{m} (\alpha_1 + \omega t) \tag{21}$$

$$\ddot{z} = -\frac{F}{m}\,(\beta_1 + \Omega t) \tag{22}$$

The terminal boundary conditions are such that the vehicle will end up on a circular orbit at the desired nominal radius. In this Cartesian co-ordinate system and with these approximate equations of motion, these boundary conditions are given by Eqs. 23-27.

$$\dot{x}_2 = \sqrt{\frac{\mu}{R}}\left[1 - \frac{\mu}{R^3}\,\frac{t_2^2}{2}\right] \tag{23}$$

$$\dot{y}_2 = -\frac{\mu}{R^2}\,t_2 \tag{24}$$

$$y_2 = -\frac{\mu}{R^2}\,\frac{t_2^2}{2} \tag{25}$$

$$\dot{z}_2 = 0 \tag{26}$$

$$z_2 = 0 \tag{27}$$

The integration of the equations of motion may be carried out by standard techniques to yield Eqs. 28-32.

$$\dot{x}_2 = \dot{x}_0 \left[1 - \frac{1}{2} \frac{\mu}{R^3} (t_2 - t_0)^2 \right] - x_0 \frac{\mu}{R^3} (t_2 - t_0) - \frac{\Delta V}{2} \left[2 - \alpha_1^2 - \beta_1^2 - \frac{\mu}{R^3} t_2^2 \right]$$

$$+ \frac{M_2}{2} \left[\frac{\mu}{R^3} + \omega^2 + \Omega^2 \right] \tag{28}$$

$$\dot{y}_2 = \dot{y}_0 - \frac{\mu}{R^2} (t_2 - t_0) - \Delta V \alpha_1 \tag{29}$$

$$y_2 = y_0 + y_0 (t_2 - t_0) - \frac{1}{2} \frac{\mu}{R^2} (t_2 - t_0)^2 - \Delta V \alpha_1 t_2 + M_2 \omega \tag{30}$$

$$\dot{z}_2 = \dot{z}_0 - \Delta V \beta_1 \tag{31}$$

$$z_2 = z_0 + \dot{z}_0 (t_2 - t_0) - \Delta V \beta_1 t_2 + M_2 \Omega \tag{32}$$

There are five quantities that must be determined in order to meet the five desired terminal conditions. These quantities are the directions of the thrust at the centroid time as well as the rates of rotation during the burn and the total duration of the burn. To this order of approximation, these quantities are given explicitly by Eqs. 33-37.

$$\alpha_1 = \frac{\dot{y}_0 + \frac{\mu}{R^2} t_0}{\Delta V} \tag{33}$$

$$\omega = - \frac{y_0 - \dot{y}_0 t_0 - \frac{\mu}{R^2} \frac{t_0^2}{2}}{M_2} \tag{34}$$

$$\beta_1 = \frac{\dot{z}_0}{\Delta V} \tag{35}$$

$$\Omega = -\frac{z_0 - \dot{z}_0 t_0}{M_2} \tag{36}$$

$$\Delta V = \left(\dot{x}_0 - \sqrt{\frac{\mu}{R}}\right)\left[1 + \frac{\alpha_1^2 + \beta_1^2}{2}\right] + \frac{\dot{x}_0}{2}\frac{\mu}{R^3}\left(2t_2 t_0 - t_0^2\right)$$

$$- x_0 \frac{\mu}{R^3}(t_2 - t_0) + \frac{M_2}{2}\left[\frac{\mu}{R^3} + \omega^2 + \Omega^2\right] \tag{37}$$

This completes the development of the equations of motion during the terminal burn as well as the solution of the boundary value problem for control of the vehicle. It is still necessary to make a feedback control law out of these equations and to determine the times at which the thrust should be turned on and turned off. An accurate way to create a feedback control law is to continually redefine the position of the centroid on the basis of the estimated time-to-go to the completion of the burn. This will cause the Cartesian co-ordinate system to rotate and change its origin during the burn but will cause the solution to become progressively more accurate as the terminal time is approached. With this type of feedback control, the engine may be turned off whenever some measure of the terminal error becomes small enough.

The optimum time to turn the engine on may be determined by noting that Robbins' theory predicts that the centroid of the final burn should be at the periapsis of the hyperbola. The required length of the burn may then be determined by substituting the control Eqs. 33-36 into Eq. 37 and referring all quantities to the perigee of the unpowered hyperbola, yielding Eq. 38. The last term in Eq. 38 is of higher order than the other terms and may generally be neglected.

$$\Delta V = \dot{x}_p - \sqrt{\frac{\mu}{R}} + \frac{\mu}{R^3}\,\frac{M_2}{2} + \frac{y_p^2 + z_p^2}{2M_2} + \frac{z_p^2}{2\Delta V} \tag{38}$$

III - Optimization of the Combined Maneuver

The first section of the analysis considered the optimum midcourse correction if there was going to be an impulsive terminal burn. The second section then developed the equations of motion and the cost for a finite thrust terminal burn. This final section is intended to determine the minimum fuel consumption for the combined maneuver with a midcourse impulse and a finite thrust terminal burn. With a finite thrust terminal burn, it is possible to produce changes normal to the nominal thrust direction during the burn, so that part of the position correction may be done during the terminal burn and part may be done during the midcourse burn. The total additional cost, over and above the ΔV of the nominal impulsive hyperbola, is given by Eq. 39.

$$\Sigma\delta V = \sqrt{1 - \left[\frac{\partial \dot{x}_p}{\partial u_{nc}}\right]^2} \cdot \sqrt{\frac{\left[y_{pi} - y_p\right]^2}{\left[\frac{\partial y_p}{\partial u_{cy}}\right]^2} + \frac{\left[z_{pi} - z_p\right]^2}{\left[\frac{\partial z_p}{\partial u_{cy}}\right]^2}}$$

$$- \frac{\partial \dot{x}_p / \partial u_{cy}}{\partial y_p / \partial u_{cy}}\,(y_{pi} - y_p) + \frac{\left(y_p^2 + z_p^2\right)}{2M_2} + \frac{\partial \dot{x}_p}{\partial y_p}\,y_p + \frac{\mu}{R^3}\,\frac{M_2}{2} \tag{39}$$

The only unknowns in Eq. 39 are y_p and z_p which are the position displacements at perigee after the midcourse impulse. All the other terms in Eq. 39 may be considered as constants and they have all been previously determined in this paper except for the one given by Eq. 40.

$$\frac{\partial \dot{x}_p}{\partial y_p} = - \frac{\mu}{R_p^2 \sqrt{e+1}} \tag{40}$$

The optimum values of y_p and z_p may then be determined by differentiating Eq. 39 with respect to these two quantities, and setting these derivatives equal to zero. This produces two simultaneous quadratic equations for y_p and z_p. These simultaneous equations may be solved by standard algebraic techniques, either analytically or by iteration. Once y_p and z_p have been determined, the components of the midcourse correction may be determined from the formulas in Section 1 by substituting the changes in these quantities for the total corrections considered in Section 1. The terminal maneuver is then carried out according to the suggestions developed in Section 2. Eq. 39 may also be used to determine the relative cost of having a midcourse correction or of correcting all the position components during the terminal burn. These comparative costs may then be used to determine the desirability of performing a midcourse correction.

CONCLUSIONS

1. The optimum midcourse guidance correction for rendezvous or orbit transfer may be found by a simple modification of the standard calculations for position guidance.

2. An accurate near-optimal terminal guidance scheme can be developed explicitly using Robbins' theory of near-impulsive transfers.

3. A unified theory of minimum-fuel guidance can be developed for a large class of missions.

POSSIBLE EXTENSIONS OF THE ANALYSIS

1. This problem can be extended to nonlinear planar corrections by using the results of Horner (Ref. 16). The adjoint solutions corresponding to small planar corrections are identical with Horner's.

2. It should be possible to generalize the analysis to arbitrary time-open maneuvers in the neighborhood of a given time-open minimum-impulse maneuver.

REFERENCES

1. Kelley, H. J.; "Guidance Theory and Extremal Fields," IRE Trans. on Automatic Control, AC-7, 5, 1962, pp. 75-82.

2. Breakwell, J. V., Speyer, J. L. and Bryson, A. E.; "Optimization and Control of Nonlinear Systems Using the Second Variation," J. SIAM Control, 1, 1963, pp. 193-223.

3. Miner, W. E., Schmieder, D. H. and Braud, N. J.; "The Path-Adaptive Mode for Guiding Space Flight Vehicles," Progress in Astronautics and Rocketry 8, Guidance and Control, Academic Press, New York, 1962.

4. Perkins, F. M.; "Explicit Tangent-Steering Guidance for Multi-Stage Boosters," Astronautica Acta, 12, 1966, pp. 212-223.

5. Chandler, D. C. and Smith, I. E.; "Development of the Iterative Guidance Mode with its Application to Various Vehicles and Missions," AIAA/JACC Guidance and Control Conference, August 15, 1966.

6. Robbins, H. M.; "Optimal Steering for Required Velocity Guidance," Navigation, 12, 1965, pp. 355-363.

7. Battin, R. H.; Astronautical Guidance, McGraw-Hill Book Co., New York, 1964.

8. Cherry, G. W.; "A General, Explicit Optimizing Guidance Law for Rocket Propelled Spaceflight," AIAA Paper 64-638, 1964.

9. Robbins, H. M.; "Analytical Study of the Impulsive Approximation," AIAA Journal, Vol. 4, No. 8, August 1966, pp. 1417-1423.

10. Noton, A. R. M., Cutting, E. and Barnes, F. L.; "Analysis of Radio-Command Midcourse Guidance," JPL Tech. Report No. 32-28, Sept. 1960.

11. Potter, J. E. and Stern, R. E.; "Optimization of Midcourse Velocity Corrections," Proc. of the IFAC Symposium on Automatic Control in Peaceful Uses of Space, Stavanger, Norway, J. A. Aseltine, ed., Plenum Press, New York, June 1965.

12. Neustadt, L. W.; "Optimization, A Moment Problem and Nonlinear Programming," Society for Industrial and Applied Mathematics (SIAM), Journal on Control, Vol. 2, No. 1, 1964, pp. 33-89.

13. Platonov, A. K.; "Investigation of the Properties of Correction Maneuvers in Interplanetary Flights," Cosmic Research, 4, 1966, pp. 587-607.

14. Lawden, D. F.; Optimal Trajectories for Space Navigation, Butterworths, London, 1963.

15. Lion, P. M. and Handelsman, M.; "The Primer Vector on Fixed-Time Impulsive Trajectories," AIAA Fifth Aerospace Sciences Meeting, New York, Paper 67-54, January 1967.

16. Horner, J. M.; "Optimum Two-Impulse Transfer Between Arbitrary Coplanar Terminals," ARS Journal, Vol. 32, No. 1, January 1962, pp. 95-96.

RESTRICTED MINIMA OF FUNCTIONS OF CONTROLS

By J. Warga
Northeastern University
Boston, Massachusetts

NASA Research Grant NGR 22-011-020

RESTRICTED MINIMA OF FUNCTIONS OF CONTROLS

1. INTRODUCTION

In a previous paper [1] we have discussed the existence of controls ρ that minimize a function x^o subject to the restrictions that, for every value of its argument t in a metric space, $\rho(t)$ is contained in some preassigned set $R^{\#}(t)$ and that $x(\rho) \in B_1$, where x is a given mapping and B_1 is a closed subset of a topological space H. We have shown that, in a large class of problems, such minimizing controls exist in a larger space of "relaxed controls" and that these relaxed controls can be approximated by original controls.

In this paper we shall assume that H is the euclidean n-space E_n. We wish to investigate certain necessary conditions for minimum that might be considered a generalization of the Weierstrass E-condition and of the transversality conditions of the calculus of variations. In this sense our results represent an extension of certain methods and theorems of the mathematical control theory, and specifically of references [2] and [3], to a more general setting. The necessary conditions that we obtain are no longer restricted, however, to minima over the space $\mathscr{S}$ of relaxed controls but apply as well to minima over the space $\mathscr{R}$ of original controls (if such minima exist). Thus our present results also generalize Pontryagin's maximum principle. Furthermore, the space $\mathscr{R}$ is no longer restricted, as in [1], to measurable mappings from one metric compact set to another.

Previous attempts to apply the methods of the mathematical control theory to problems involving functions defined otherwise than by a system of differential or difference equations were mostly limited to special, and linear, problems. Recent results of Neustadt [4, 5] are, however, quite general. They are based on a separation theorem for convex sets that represent certain linearizations of constraints. Our approach is, however, different from Neustadt's; in particular, our basic results are stated directly in the form of inequalitites involving the value of the minimizing control at an arbitrary point rather than in the form of functional inequalities.

Let T and R be arbitrary sets, B a convex set, $\mathscr{R}$ a class of controls, that is, mappings from T to R, $x = (x^1, \ldots, x^n)$ a given function from $\mathscr{R} \times B$ to E_n, and B_1 a given set in E_n. We wish to characterize a control $\bar{\rho} \in \mathscr{R}$ and a point $\bar{b} \in B$ that yield a minimum of $x^1(\rho, b)$ subject to the condition $x(\bar{\rho}, b) \in B_1$. The necessary conditions for minimum that we derive are expressed in terms of certain variational derivatives $Dx(\bar{\rho}, \bar{b}; t^*, \rho^*)$ respectively $Dx(\bar{\rho}, \bar{b}; t^*, r)$ defined in section 2. These derivatives represent, roughly speaking, the rate of change of x when its argument $\bar{\rho}$ is replaced by the function ρ^* (respectively the constant function r) over a "small" set in the "neighborhood" of t^*.

As an illustration, we consider, in sections 3 and 5, the "standard" problem of the mathematical control theory of ordinary differential equations and prove a slight generalization of the usual necessary conditions.

2. NECESSARY CONDITIONS FOR MINIMUM

Let T and R be arbitrary sets, B a convex set, $\mathscr{R}$ an arbitrary class of mappings from T to R, and B_1 a set in E_n. The vector function $x = (x^1, \ldots, x^n)$ is a given mapping from $\mathscr{R} \times B$ to E_n. If $\rho : T \to R$, we denote by $\rho(t)$ the image, under the mapping ρ, of a point t in T. If the mapping ρ depends on some parameters a, b, c, we designate by $\rho(a, b, c)$, or by $\rho(\cdot\, ; a, b, c)$ the mapping, and by $\rho(t; a, b, c)$ the image of t under the mapping. Similarly x denotes a mapping and $x(\rho)$ the image of ρ under the mapping x. We also write, when it appears more appropriate, $t \to \rho(t)$ to represent a mapping.

If ρ_1 and ρ_2 are two mappings from T to R, and A is a subset of T, we designate by $[\rho_1, A; \rho_2]$ the mapping ρ defined by the relations

$$\rho(t) = \rho_1(t) \text{ on } A, \quad \rho(t) = \rho_2(t) \text{ on } T\text{-}A.$$

Similarly, if $\rho_1, \rho_2, \ldots, \rho_k, \bar{\rho}$ are mappings from T to R, and $A_1, A_2, \ldots, A_k$ are disjoint subsets of T, we designate by $[\rho_1, A_i(i=1,\ldots,k); \bar{\rho}]$ the mapping ρ defined by the relations

$$\rho(t) = \rho_i(t) \text{ on } A_i(i=1,\ldots,k),\ \rho(t) = \bar{\rho}(t) \quad \text{elsewhere on T.}$$

Let T* be a subset of T, and let $\mathscr{M}$ be a collection of subsets $M(t,\alpha)$ of $T(t \in T^*,\ \alpha \geqslant 0)$. Let $\bar{\rho} \in \mathscr{R}$, $\rho^* \in \mathscr{R}$, $\bar{b} \in B$, $t^* \in T$, $\alpha > 0$, and let $\rho' = [\rho^*, M(t^*,\alpha); \bar{\rho}]$.

If $\rho' \in \mathscr{R}$ for sufficiently small α, and if

$$\lim_{\alpha \to +0} \frac{1}{\alpha} (x(\rho', \bar{b}) - x(\bar{\rho}, \bar{b}))$$

exists, we shall say that "x has an $\mathscr{M}$-derivative at $(\bar{\rho}, \bar{b})$ with respect to (t^*, ρ^*)" and we shall designate this limit by $D_{\mathscr{M}}\, x(\bar{\rho}, \bar{b}; t^*, \rho^*)$. If $\mathscr{R}^*(t^*)$ is a subset of $\mathscr{R}$ for each $t^* \in T^*$ and $D_{\mathscr{M}}\, x(\bar{\rho}, \bar{b}; t^*, \rho^*)$ is the same for all $\rho^* \in \mathscr{R}^*(t^*)$ such that $\rho^*(t^*) = r$, we shall write

$$D_{\mathscr{M};\mathscr{R}^*}\, x(\bar{\rho}, \bar{b}; t^*, r) \quad \text{or} \quad Dx(\bar{\rho}, \bar{b}; t^*, r)$$

(if $\mathscr{M}$ and the mapping $t^* \to \mathscr{R}^*(t^*)$ are fixed).

Let now $\bar{\rho} \in \mathscr{R}$, $\bar{b} \in B$, and $b \in B$. We shall write $Dx(\bar{\rho}, \bar{b}; b)$ for $\lim_{\theta \to +0} \frac{1}{\theta} (x(\bar{\rho}, (1-\theta)\bar{b} + \theta b) - x(\bar{\rho}, \bar{b}))$.

<u>Definition 2.1. Local variations in $\mathscr{R} \times B$.</u>

Let $\bar{\rho} \in \mathscr{R}$, $\bar{b} \in B$, $T^* \subset T$, let $\mathscr{N}_k$ be, for $k=1,2,\ldots,n^2$, a collection of subsets $N_k(t,\alpha)$ of $T(t \in T^*,\ \alpha \geqslant 0)$, and let $\mathscr{N} = \left\{ \mathscr{N}_k \mid k=1,\ldots,n^2 \right\}$. Let $\mathscr{R}^*$ be a mapping from T* to the class of subsets of $\mathscr{R}$.

We shall say that $(T^*, \mathscr{R}^*, \mathscr{N})$ define "local variations for x in $\mathscr{R} \times B$ at $(\bar{\rho}, \bar{b})$" if the following conditions hold:

(2.1.1) For $t^*, t_1^*, t_2^* \in T^*$, $k, k_1, k_2 = 1, 2, \ldots, n^2$, and $\alpha, \beta \geqslant 0$,

$$N_k(t^*, \alpha) \subset N_k(t^*, \beta) \quad \text{if} \quad \alpha \leqslant \beta; \; N_k(t^*, 0)$$

is the empty set; $N_{k_1}(t^*, \alpha)$ and $N_{k_2}(t^*, \beta)$ are disjoint if $k_1 \neq k_2$; and $N_{k_1}(t_1^*, \alpha)$ and $N_{k_2}(t_2^*, \beta)$ are disjoint if $t_1^* \neq t_2^*$ and both α and β are sufficiently small.

(2.1.2) Let an array with elements $\beta^{ij}(i, j = 1, \ldots, n)$ be represented by $\beta^{\square}$. Let, for every choice of $t^{\square}$ with elements $t^{ij} \in T^*$ and of $\rho^{\square}$ with elements $\rho^{ij} \in \mathscr{R}^*(t^{ij})$, the set $\Omega = \Omega(t^{\square})$ in E_{n^2} contain all arrays $\omega^{\square}$ with $\omega^{ij} \geqslant 0$ which are such that the sets $N_{nj-n+i}(t^{ij}, \omega^{ij})$ $(i, j = 1, \ldots, n)$ are disjoint, and let $\rho' = \rho'(t^{\square}, \rho^{\square}, \omega^{\square}) = [\rho^{ij}, N_{nj-n+i}(t^{ij}, \omega^{ij})(i, j = 1, \ldots, n); \bar{\rho}]$ for $\omega^{\square} \in \Omega(t^{\square})$. Finally let $b^{\square}$ have elements $b^{ij} \in B$, $0^{\square}$ have elements

$$0^{ij} = 0, \; \mathscr{T} = \left\{ \theta^{\square} \mid \theta^{ij} \geqslant 0, \sum_{i,j=1}^{n} \theta^{ij} \leqslant 1 \right\}, \; \theta^{o} = 1 - \sum_{i,j=1}^{n} \theta^{ij},$$

and

$$\theta^{\square} \cdot b^{\square} = \theta^{o} \bar{b} + \sum_{i,j=1}^{n} \theta^{ij} b^{ij} .$$

Then

(2.1.2.1) $\rho' \in \mathscr{R}$;

(2.1.2.2) for fixed $t^{\square}$, $\rho^{\square}$ and $b^{\square}$, the function $(\omega^{\square}, \theta^{\square}) \to \xi(\omega^{\square}, \theta^{\square}) = \xi(\omega^{\square}, \theta^{\square}, t^{\square}, \rho^{\square}, b^{\square}) = x(\rho'(t^{\square}, \rho^{\square}, \omega^{\square}), \theta^{\square} \cdot b^{\square})$ from $\Omega \times \mathscr{T}$ to E_n is continuous in some neighborhood of $(0^{\square}, 0^{\square})$, and has a differential at $(0^{\square}, 0^{\square})$ (relative to $\Omega \times \mathscr{T}$);

(2.1.2.3) for every $t^* \in T^*$, $\rho^* \in \mathscr{R}^*(t^*)$, and $k=1,2,\ldots,n^2$, $D_{\mathscr{N}_k} x(\bar{\rho},\bar{b};t^*,\rho^*) = D\, x(\bar{\rho},\bar{b};t^*,r)$ exists, is independent of k, and has the same value for all $\rho^* \in \mathscr{R}^*(t^*)$ such that $\rho^*(t^*) = r$.

We can now state our general necessary conditions for minimum which we shall prove in section 4.

Theorem 2.2. Let $(\bar{\rho},\bar{b})$ yield the minimum of $x^1(\rho,b)$ in $\mathscr{R} \times B$ subject to the condition $x(\rho,b) \in B_1$. Let $(T^*,\mathscr{R}^*,\mathscr{N})$ define local variations for x in $\mathscr{R} \times B$ at $(\bar{\rho},\bar{b})$, and let, for all $t^* \in T^*$, $R^*(t^*) = \{\rho^*(t^*) \mid \rho^* \in \mathscr{R}^*(t^*)\}$. Let, furthermore, B_1^* be a convex set in E_m, $\bar{b}_1^*$ a point in B_1^*, and $\phi: B_1^* \to B_1$ a continuous mapping such that $\phi(\bar{b}_1^*) = x(\bar{\rho},\bar{b})$, $\phi(B_1^*) \subset B_1$ and ϕ has a differential at $\bar{b}_1^*$ (relative to B_1^*) $d\phi(\bar{b}_1^*;b_1^* - \bar{b}_1^*) = \phi_{b_1^*}(\bar{b}_1^*)(b_1^* - \bar{b}_1^*)$ (where $\phi_{b_1^*}(\bar{b}_1^*)$ is a linear operator from E_m to E_n). Then either

$$(2.2.1) \qquad \phi^1_{b_1^*}(\bar{b}_1^*)\bar{b}_1^* = \operatorname*{Min}_{b_1^* \in B_1^*} \phi^1_{b_1^*}(\bar{b}_1^*)b_1^*,$$

or there exists a nonvanishing vector λ in E_n such that

$$(2.2.2) \qquad \lambda \cdot Dx(\bar{\rho},\bar{b};t^*,r) \geqslant 0 \text{ for all } t^* \in T^* \text{ and } r \in R^*(t^*);$$

$$(2.2.3) \qquad \lambda \cdot Dx(\bar{\rho},\bar{b};b) \geqslant 0 \text{ for all } b \in B;$$

and

$$(2.2.4) \quad (\mu^o\delta_1 - \lambda) \cdot \left[\phi_{b_1^*}(\bar{b}_1^*)\bar{b}_1^*\right] = \operatorname*{Min}_{b_1^* \in B_1^*} (\mu^o\delta_1 - \lambda) \cdot \left[\phi_{b_1^*}(\bar{b}_1^*)b_1^*\right]$$

for some $\mu^o \geqslant 0$, where $\delta_1 = (1,0,\ldots,0) \in E_n$.

Remark. Relation (2.2.2) generalizes the Weierstrass E-condition, relation (2.2.3) generalizes the transversality conditions at the initial point and describes the dependence on parameters, and relation (2.2.4) generalizes the transversality conditions at the endpoint.

Theorem 2.2 is of particular interest in the case [1] when $t \to R^{\#}(t) \subset R$ is a given mapping from T to the class of nonempty subsets of R, and $\mathscr{R}$ is the class of measurable relaxed controls σ such that the probability measure $\sigma(t)$ is supported on the closure of $R^{\#}(t)$ for all $t \in T$. We may then assert [1, Th. 2.6] that in a large class of problems there exists a relaxed control $\bar{\sigma}$ and a point $\bar{b}$ that yield the restricted minimum assumed in Theorem 2.2; and we may verify a priori the other assumptions of Theorem 2.2. We are then able to state that a minimizing control $\bar{\sigma}$ and point $\bar{b}$ exist and either satisfy condition (2.2.1) or conditions (2.2.2), (2.2.3), and (2.2.4). Since these relations often admit only a finite number of solutions, we can determine a minimizing $\bar{\sigma}$ and $\bar{b}$; in this sense, [1, Th. 2.6] and Theorem 2.2 often provide constructive conditions for minimum.

3. FUNCTIONS OF CONTROLS DEFINED BY ORDINARY DIFFERENTIAL EQUATIONS

We shall now illustrate the use of Theorem 2.2 in certain standard problems of the control theory, postponing the proof of the results presented in this section to section 5. Let T be the closed interval $[t_0, t_1]$ of the real axis, R a separable metric space, $R^{\#}$ a mapping from T to the class of nonempty subsets of R, $B_0 \subset E_n$, $B_1 \subset E_n$, and $g: E_n \times T \times R \to E_n$. In this section, and in section 5, the words <u>measure</u> and <u>measurable</u> will be used in the sense of Lebesgue and $|A|$ will represent the measure of $A \subset T$.

Let $\mathscr{R}'$ be a class of mappings $\rho: T \to R$ such that $t \to g(v, t, \rho(t))$ is measurable on T for every $v \in E_n$ and $\rho \in \mathscr{R}'$ and $[\rho_i, A_i (i=1, \dots, k); \rho] \in \mathscr{R}'$ if k is a positive integer, each A_i is a denumerable union of intervals, and $\rho \in \mathscr{R}'$, $\rho_i \in \mathscr{R}'$ $(i=1, \dots, k)$. We shall henceforth refer to elements of $\mathscr{R}'$ as "measurable" mappings (as distinguished from measurable mappings). We set $\mathscr{R} = \{\rho \in \mathscr{R}' \mid \rho(t) \in R^{\#}(t) \text{ on } T\}$.

For $\rho \in \mathscr{R}$ and $b_o \in B_o$, we consider an absolutely continuous function $y = y(\cdot\ ; \rho, b_o)$ on T such that

(3.1.1) $\quad dy(t)/dt = \dot{y}(t) = g(y(t), t, \rho(t))$ a.e. in T

and

(3.1.2) $\quad y(t_o) = b_o$.

We wish to investigate certain properties of a point $\bar{b}_o \in B_o$ and a mapping $\bar{\rho} \in \mathscr{R}$ that minimize $y^1(t_1; \rho, b_o)$ on $\mathscr{R} \times B_o$ subject to the restriction that $y(t_1; \rho, b_o) \in B_1$.

We shall say that a sequence $\{M_j\}_{j=1}^{\infty}$ of closed subsets of T is "regular at $\bar{t}$" if $|M_j| \to 0$ as $j \to \infty$, $\bar{t} \in M_j$, and diameter $(M_j) < c|M_j|$ for some positive c and all $j = 1, 2, \ldots$. We shall say that a "measurable" mapping $\rho^*: T \to R$ is "admissible at $\bar{t}$" if $\rho^*(t) \in R^{\#}(t)$ on T and $\lim_{j \to \infty} \frac{1}{|M_j|} \int_{M_j} g(v, t, \rho^*(t))\, dt = g(v, \bar{t}, \rho^*(\bar{t}))$ for all $v \in E_n$ and all sequences $\{M_j\}$ that are regular at $\bar{t}$.

We set

$$R^*(\bar{t}) = \{\rho^* \mid \rho^* \text{ is admissible at } \bar{t}\}$$

and

$$R^*(\bar{t}) = \{\rho^*(\bar{t}) \mid \rho^* \in \mathscr{R}^*(\bar{t})\} \quad .$$

We shall also write $\overline{R^*(\bar{t})}$, $\bar{M}_j$, etc. to represent closures of the sets $R^*(\bar{t})$, M_j, etc.

Assumption 3.2. For every $v \in E_n$, $r \in R$, and $t \in T$, the function $g(v, t, \cdot)$ is continuous on R, $g(v, \cdot, r)$ is measurable on T, and $g(\cdot, t, r)$ is continuous and has continuous first order derivatives on E_n. Furthermore, for every v in E_n there exists an integrable function ψ_v on T such that $|g(v, t, r)| \leq \psi_v(t)$ on T x R. Finally, for every bounded subset D of E_n there exists an integrable function ψ_D on T such that $|g_v(v,t,r)| \leq \psi_D(t)$ on D x T x R. Here $|g| = \left(\sum_{j=1}^{n} (g^j)^2\right)^{1/2}$, g_v is the matrix $(\partial g^i/\partial v^j)$, and $|g_v| = \sum_{i,j=1}^{n} |\partial g^i/\partial v^j|$.

Remark. Assumption 3.2 implies that we may choose as $\mathcal{R}'$ the class of all the measurable functions from T to R.

Theorem 3.3. Let $(\bar{\rho}, \bar{b}_o)$ minimize $y^1(t_1;\rho, b_o)$ among all points b_o in B_o and all "measurable" mappings ρ such that $\rho(t) \in R^{\#}(t)$ on T and $y(t_1;\rho, b_o) \in B_1$, and let Assumption 3.2 be satisfied. Let $\bar{y} = y(\cdot\,;\bar{\rho}, \bar{b}_o)$, $\bar{b}_1 = \bar{y}(t_1)$ and let, for k = 0, 1, B^*_k be a convex set in E_{m_k} and $\phi_k = (\phi^1_k, \ldots, \phi^n_k) : E_{m_k} \to E_n$ a continuously differentiable mapping such that $\phi_k(B^*_k) \subset B_k$ and $\phi_k(\bar{b}^*_k) = \bar{b}_k$ for some $\bar{b}^*_k \in B^*_k$. Let A_k be, for k = 0, 1, the matrix $(\partial\phi^i_k/\partial b^{*j}_k)$ evaluated at $\bar{b}^*_k$, and let A^i_k be the i-th row of this matrix. Then either

$$(3.3.1)\qquad A^1_1 \cdot \bar{b}^*_1 = \min_{b^*_1 \in B^*_1} A^1_1 \cdot b^*_1 ,$$

or there exists an absolutely continuous function $z:T \to E_n$ such that

$$(3.3.2)\qquad \dot{\bar{y}}(t) = g(\bar{y}(t), t, \bar{\rho}(t)) \quad \text{a.e. in T,}$$

$$(3.3.3)\qquad \dot{z}(t) = -\,g^T_v(\bar{y}(t), t, \bar{\rho}(t))\, z(t) \quad \text{a.e. in T}$$

(where g^T_v is the transpose of the matrix g_v),

$$(3.3.4)\qquad |z(t)| \neq 0 \text{ on T,}$$

$$(3.3.5)\qquad z(t)\cdot g(\bar{y}(t), t, \bar{\rho}(t)) = \min_{r \in \bar{R}^{\#}(t)} z(t)\cdot g(\bar{y}(t), t, r) \quad \text{a.e. in T,}$$

$$(3.3.6)\qquad z(t_o)\cdot A_o \bar{b}^*_o = \min_{b^*_o \in B^*_o} z(t_o)\cdot A_o b^*_o ,$$

and

$$(3.3.7) \qquad (\gamma\delta_1 - z(t_1)) \cdot A_1\bar{b}_1^* = \min_{b_1^* \in B_1^*} (\gamma\delta_1 - z(t_1)) \cdot A_1 b_1^*$$

for some $\gamma \geq 0$. Hence $\delta_1 \geq (1, 0, \ldots, 0) \in E_n$.

In particular, if $R^{\#}(t) = R$ on T and $\mathscr{R}'$ contains the constant mapping $t \to r$ for all r in a dense denumerable subset R_∞ of R, then $\bar{R}^*(t)$ can be replaced by R in statement (3.3.5).

By combining [1, Th. 3.1] and Theorem 3.3, we can prove the existence of a minimizing relaxed control $\bar{\sigma}$ and a point $\bar{b}_o$ and can state some of their characteristic properties. We require

<u>Assumption 3.4.</u>

(3.4.1) $\quad R$ is compact.

(3.4.2) $\quad B_o$ is compact and B_1 is closed.

(3.4.3) $\quad$ There exists an integrable function ψ on T such that $|g_v(v, t, r)| \leq \psi(t)$ on $E_n \times T \times R$.

(3.4.4) $\quad$ Let $\mathscr{R}^{\#} = \{\rho: T \to R \mid \rho(t) \in R^{\#}(t)$ on T and ρ is measurable$\}$. Then for every ϵ there exists a closed subset T_ϵ of T, of measure at least $|T| - \epsilon$, with the property that (a) for every $\bar{t} \in T_\epsilon$ and every $r \in \overline{R^{\#}(\bar{t})}$ there exists a mapping $\rho \in R^{\#}$, continuous at $\bar{t}$ when restricted to T_ϵ, and such that the distance from $\rho(\bar{t})$ to r is at most ϵ; and (b) for every $\bar{t} \in T_\epsilon$ and every $h > 0$ there exists a positive $\delta = \delta(h, \bar{t})$ such that $R^{\#}(t)$ and $P^{\#}(\bar{t})$ are in the h-neighborhood of each other if $t \in T_\epsilon$ and $|t - \bar{t}| \leq \delta$.

Now let S be the class of regular Borel probability measures on R. It is well known [9, p. 426] that a metric can be defined on S such that S is separable and the convergence in S is the weak convergence of measures: that is, a sequence $s_1, s_2, \ldots$ converges to s in S if $\int_R c(r)s_j(dr) \to \int_R c(r)s(dr)$ as $j \to \infty$ for every continuous

$c:R \to E_1$. Let $\mathscr{A}^{\#}$ be the set of mappings σ from T to S such that $\sigma(\bar{R}^{\#}(t);t) = 1$ on T and $t \to \int_R c(r)\sigma(dr;t)$ is measurable on T for every continuous $c:R \to E_1$. Here $\sigma(R';t)$ is the $\sigma(t)$-measure of a subset $R' \subset R$.

We refer to an absolutely continuous function $\xi:T \to E_n$ as a "relaxed curve" if $\xi(t_o) \in B_o$ and $\dot{\xi}(t)$ belongs, a.e. in T, to the convex closure of the set $\{g(\xi(t), t, r) \mid r \in R^{\#}(t)\}$. This definition is equivalent, in view of our assumptions and of [1, Th. 3.1], to the statement that $\xi = \xi(\cdot\,;\sigma, b_o)$ satisfies the relations

$$\dot{\xi}(t) = \int_R g(\xi(t), t, r)\,\sigma(dr;t) \quad \text{a.e. in T,}$$

$$\xi(t_o) = b_o,$$

for some $\sigma \in \mathscr{A}^{\#}$ and $b_o \in B_o$. (This definition is also consistent with the one in [1, section 3] for $A = E_n$.)

We can state

<u>Theorem 3.5.</u> Let $B_o, B_1, T, R, R^{\#}$, and g satisfy Assumptions 3.2 and 3.4, and assume that $\xi(t_1;\sigma', b_o') \in B_1$ for some $\sigma' \in \mathscr{A}^{\#}$ and $b_o' \in B_o$. Then there exist a relaxed control $\bar{\sigma}$ and a point $\bar{b}_o \in B_o$ that minimize $\xi^1(t_1;\sigma, b_o)$ on $\mathscr{A}^{\#} \times B_o$ subject to the condition that $\xi(t_1;\sigma, b_o) \in B_1$; and the corresponding minimizing relaxed curve $\bar{\xi} = \xi(\cdot\,;\bar{\sigma}, \bar{b}_o)$ can be uniformly approximated on T by a sequence $\xi_1, \xi_2, \ldots$ of absolutely continuous curves such that $\dot{\xi}_j(t) = g(\xi_j(t), t, \rho_j(t))$ a.e. in T $(j = 1, 2, \ldots)$, the mappings ρ_j are measurable, and $\rho_j(t) \in R^{\#}(t)$ on T.

Let $f(v, t, s) = \int_R g(v, t, r)s(dr)$ on $E_n \times T \times S$, let $S^*(t) = \{s \in S \mid s(\bar{R}^*(t)) = 1\}$ for $t \in T$, let $\bar{b}_1 = \xi(t_1;\bar{\sigma}, \bar{b}_o)$, and let E_{m_k}, B_k^*, ϕ_k, and A_k be defined as in the statement of Theorem 3.3. Then either condition (3.3.1) or conditions (3.3.2) through (3.3.7) of Theorem 3.3 are satisfied, with $\bar{y}, g, \bar{\rho}, R^*$, and r replaced by, respectively, $\bar{\xi}, f, \bar{\sigma}, S^*$, and s.

Furthermore, condition (3.3.5) of Theorem 3.3 implies that, a.e. in T,

$$(3.5.1)\qquad z(t)\cdot g(\bar{\xi}(t),t,\bar{r}) = \operatorname*{Min}_{r\in R}\ z(t)\cdot g(\bar{\xi}(t),t,r)$$

for every $\bar{r}$ in the support of $\bar{\sigma}(t)$, if $R^{\#}(t) = R$ on T.

4. PROOF OF THEOREM 2.2

The proof of Theorem 2.2 is essentially contained in the lemma that follows and that resembles, in many respects, Lemma 3.1 of [3, p. 132]. The convex set W is patterned after a construction of McShane [6, pp. 17-18]. Brouwer's fixed point theorem appears to have been first applied in a similar context by H. Halkin [7, p. 75].

Lemma 4.1. Let $(\bar{\rho},\bar{b})$ minimize $x^1(\rho,b)$ in $\mathscr{R}$ x B subject to the conditions $x^{\ell}(\rho,b) = 0(\ell=2,\ldots,n)$. Let $T^*,\mathscr{R}^*,\mathscr{N}$ define local variations for x in $\mathscr{R}$ x B at $(\bar{\rho},\bar{b})$. Then there exists a nonvanishing vector λ in E_n such that $\lambda^1 \geqslant 0$,

$$(4.1.1)\qquad \lambda\cdot Dx(\bar{\rho},\bar{b};t^*,r) \geqslant 0 \text{ for all } t^*\in T^* \text{ and } r\in R^*(t^*),$$

and

$$(4.1.2)\qquad \lambda\cdot Dx(\bar{\rho},\bar{b};b) \geqslant 0 \text{ for all } b\in B.$$

Proof. We shall use the notation that we have introduced in section 2. Let $V_1 = \{Dx(\bar{\rho},\bar{b};t^*,r)\,|\,t^*\in T^*,\ r\in R^*(t^*)\}$, $V_2 = \{Dx(\bar{\rho},\bar{b};b)\,|\,b\in B\}$, and let W be the convex cone in E_n generated by $V_1\cup V_2$; that is

$$W = \{a^1v_1+\ldots+a^nv_n\,|\,v_i\in V_1\cup V_2,\ a^i \geqslant 0(i=1,\ldots,n)\}.$$

Assume now, by way of contradiction, that there exists no vector λ with the stated properties. Then we can easily deduce from elementary properties of convex sets that there exists a point $w = (w^1,0,\ldots,0)$ in the interior of W, linearly independent vectors (points) $w_i \in W$, and positive numbers $c^i (i=1,\ldots,n)$ such that

$$(4.1.3) \qquad W^1 < 0 \text{ and } w = \sum_{i=1}^{n} c^i w_i .$$

By the definition of W, there exist points $t^{ij} \in T^*$, controls $\rho^{ij} \in \mathscr{R}^*(t^{ij})$, points $b^{ij} \in B$ and numbers $a^{ij} \geq 0$ $(i,j=1,\ldots,n)$ such that

$$(4.1.4) \qquad w_i = \sum_{j=1}^{n} a^{ij} \cdot Dx(\bar{\rho},\bar{b};t^{ij},\rho^{ij},b^{ij}) \quad (i=1,\ldots,n)$$

where, for each i, j, $Dx(\bar{\rho},\bar{b};t^{ij},\rho^{ij},b^{ij})$ either represents $Dx(\bar{\rho},\bar{b};t^{ij},\rho^{ij})$ (and is independent of b^{ij}), or represents $Dx(\bar{\rho},\bar{b};b^{ij})$ (and is independent of t^{ij} and ρ^{ij}). The matrix (w_i^{ℓ}) $(i,\ell=1,\ldots,n)$ is nonsingular since the vectors w_i are linearly independent.

Let now $\bar{\alpha}$ be sufficiently small so that the sets $N_{k_1}(t^{i_1 j_1},\alpha)$ and $N_{k_2}(t^{i_2 j_2},\beta)$ are disjoint if $(k_1, t^{i_1 j_1}) \neq (k_2, t^{i_2 j_2})$, $\alpha \leq \bar{\alpha}$ and $\beta \leq \bar{\alpha}$, let $\bar{\alpha} < 1$, and let

$$\Delta = \left\{ \delta \in E_n \mid 0 \leq \delta^i \leq \bar{\alpha} \Big/ \left(\sum_{k,j=1}^{n} a^{kj} \right) (i=1,\ldots,n) \right\} .$$

For every $\delta \in \Delta$, let $\omega^{ij}(\delta) = a^{ij}\delta^i$ and $\theta^{ij}(\delta) = 0$ respectively $\omega^{ij}(\delta) = 0$ and $\theta^{ij}(\delta) = a^{ij}\delta^i$ if $Dx(\rho,\bar{b};t^{ij},\omega^{ij},b^{ij})$ represents $Dx(\bar{\rho},\bar{b};t^{ij},\rho^{ij})$ respectively $Dx(\rho,b;b^{ij})$. We observe that the sets $N_{nj+i-n}(t^{ij},\omega^{ij}(\delta))$ are disjoint and

$$\sum_{i,j=1}^{n} \theta^{ij}(\delta) \leq 1 \text{ for } \delta \in \Delta .$$

We now consider, for each $\delta \in \Delta$, the "perturbed" mapping $\rho'(\delta) = \rho'(t^{\square}, \rho^{\square}, \omega^{\square}(\delta)) = [\rho^{ij}, N_{nj-n+i}(t^{ij},\omega^{ij}(\delta))\ (i,j=1,\ldots,n); \bar{\rho}]$ in $\mathscr{R}$ and the "perturbed" point $b'(\delta) = \theta^{\square}(\delta) \cdot b^{\square}$ in B. By condition (2.1.2.2), the function $\delta \to \tilde{\xi}(\delta) = \xi(\omega^{\square}(\delta), \theta^{\square}(\delta)) = x(\rho'(\delta), b'(\delta))$ from Δ to E_n is continuous in some neighborhood Δ' (relative to Δ) of the origin 0 of E_n and has a differential at 0 (relative to Δ). Furthermore, by (4.1.4), the right-hand derivative

$$\partial\tilde{\xi}(\delta)/\partial\delta^k\Big|_{\delta=0} = \sum_{i,j=1}^{n} (\partial\xi(\omega^{\square}(\delta),\theta^{\square}(\delta))/\partial\omega^{ij}\cdot\partial\omega^{ij}(\delta)/\partial\delta^k + \partial\xi(\omega^{\square}(\delta),$$

$$\theta^{\square}(\delta))/\partial\theta^{ij}\cdot\partial\theta^{ij}(\delta)/\partial\delta^k)\Big|_{\delta=0} = \sum_{j=1}^{n} Dx(\bar{\rho},\bar{b};t^{kj},\rho^{kj},b^{kj})\,a^{kj} = w_k \ (k=1,\ldots,n).$$

Thus the jacobian matrix $\tilde{\xi}_\delta(0) = (\partial\tilde{\xi}^i(0)/\partial\delta^j)_{i,j=1}^n$ is nonsingular.

Let $a(\delta) = (\tilde{\xi}_\delta(0))^{-1}(\tilde{\xi}(\delta) - \tilde{\xi}(0) - \tilde{\xi}_\delta(0)\cdot\delta)$, and let $c = (c^1,\ldots,c^n)$.
We shall now show that the equation

$$(4.1.5) \qquad \delta = \gamma c - a(\delta)$$

has a solution $\delta(\gamma)$ for all sufficiently small positive γ. Indeed, since $\tilde{\xi}(\cdot)$ has a differential at 0 (relative to Δ), there exists a positive β_o such that

$$|a^i(\delta)| \le \frac{1}{4}\,\frac{c_{\min}}{c_{\max}}\,\delta_{\max}$$

and

$$\delta\in\Delta' \quad\text{if}\quad 0\le\delta^i\le\beta_o (i=1,\ldots,n),$$

where

$$c_{\min} = \operatorname*{Min}_i c^i,\ c_{\max} = \operatorname*{Max}_i c^i,\ \text{and}\ \delta_{\max} = \operatorname*{Max}_i \delta^i.$$

Let

$$0<\beta\le\beta_o,\ \gamma = \frac{1}{2}\,\beta/c_{\max},$$

and

$$\Delta''_\gamma = \left\{\delta\in\Delta' \;\middle|\; |\delta^i - \gamma c^i| \le \frac{1}{2}\gamma c_{\min}\ (i=1,2,\ldots,n)\right\}.$$

Then we can easily verify that Δ''_γ is homeomorphic to a closed ball in E_n, and

$\gamma c - a(\delta)$ is a continuous mapping of Δ''_γ into itself. Thus, by Brouwer's fixed point theorem, there exists $\delta = \delta(\gamma)$ satisfying equation (4.1.5). It follows then from relations (4.1.3) and (4.1.5) that

$$\tilde{\xi}(\delta(\gamma)) - \tilde{\xi}(0) = \gamma\tilde{\xi}_\delta(0)\, c = \gamma \sum_{i=1}^{n} c^i w_i = \gamma w = (\gamma w^1, 0, \ldots, 0);$$

hence

$$\tilde{\xi}^1(\delta(\gamma)) = \tilde{\xi}^1(0) + \gamma w^1 < \tilde{\xi}^1(0) = x^1(\bar{\rho}, \bar{b})$$

and

$$\tilde{\xi}^\ell(\delta(\gamma)) = 0 \quad (\ell = 2, \ldots, n).$$

Since $\rho'(\delta(\gamma)) \in \mathscr{R}$ and $b'(\delta(\gamma)) \in B$ for all $\beta \leq \beta_0$ and $\delta \in \Delta''_\gamma \subset \Delta$ and since $\delta(\gamma) \in \Delta''_\gamma$ and $\tilde{\xi}(\delta(\gamma)) = x(\rho'(\delta(\gamma)), b'(\delta(\gamma)))$, we conclude that, contrary to assumption, $(\bar{\rho}, \bar{b})$ does not minimize $x^1(\rho, b)$ subject to the restriction that $x^\ell(\rho, b) = 0$ $(\ell = 2, \ldots, n)$. This completes the proof of the lemma.

4.2 <u>Proof of Theorem 2.2.</u> Let $c = (b, b_1^*)$ for $b \in B$ and $b_1^* \in B_1^*$, let $\bar{c} = (\bar{b}, \bar{b}_1^*)$, and let $C = B \times B_1^*$. Then $(\bar{\rho}, \bar{c})$ minimizes $x^1(\rho, b)$ on $\mathscr{R} \times C$ subject to the restrictions $x^\ell(\rho, b) - \phi^\ell(b_1^*) = 0$ $(\ell = 1, 2, \ldots, n)$.

Let the function $y = (y^0, y^1, \ldots, y^n)$ on $\mathscr{R} \times C$ be defined by

$$y^0(\rho, c) = y^0(\rho, b, b_1^*) = x^1(\rho, b),$$

$$y^\ell(\rho, c) = y^\ell(\rho, b, b_1^*) = x^\ell(\rho, b) - \phi^\ell(b_1^*) \qquad (\ell = 1, \ldots, n).$$

Then we verify that $(T^*, \mathscr{R}^*, \mathscr{N})$ define local variations for y in $\mathscr{R} \times C$ at $(\bar{\rho}, \bar{c})$. It follows then, by Lemma 4.1, that there exists a nonvanishing vector $\mu = (\mu^0, \mu^1, \ldots, \mu^n)$

in E_{n+1} and a vector $\lambda = (\mu^0 + \mu^1, \mu^2, \ldots, \mu^n)$ in E_n such that $\mu^0 \geqslant 0$ and

(4.2.1) $\quad \mu \cdot Dy(\bar{\rho}, \bar{c}; t^*, r) = \lambda \cdot Dx(\bar{\rho}, \bar{b}; t^*, r) \geqslant 0$ for all $t^* \in T^*$ and $r \in R^*(t^*$

and

(4.2.2) $\quad \mu \cdot Dy(\bar{\rho}, \bar{c}; c) = \mu^0 Dx^1(\bar{\rho}, \bar{b}; b) + (\lambda - \mu^0 \delta_1) \cdot (Dx(\bar{\rho}, \bar{b}; b) -$

$$\phi_{b_1^*}(\bar{b}_1^*)(b_1^* - \bar{b}_1^*)) \geqslant 0 \quad \text{for all } b \in B$$

and $b_1^* \in B_1^*$, where $\delta_1 = (1, 0, \ldots, 0) \in E_n$. We observe that $Dx(\bar{\rho}, \bar{b}; \bar{b}) = 0$; hence, setting $b = \bar{b}$ in (4.2.2), it follows that

$$(\lambda - \mu^0 \delta_1) \cdot \phi_{b_1^*}(\bar{b}_1^*)(b_1^* - \bar{b}_1^*) \leqslant 0 \quad \text{for all } b_1^* \in B_1^*.$$

Since μ is nonvanishing, either λ is nonvanishing or $\mu = (\mu^0, -\mu^0, 0, \ldots, 0)$, $\mu^0 > 0$, and

$$\phi^1_{b_1^*}(\bar{b}_1^*)\bar{b}_1^* = \min_{b_1^* \in B_1^*} \phi^1_{b_1^*}(\bar{b}_1^*) b_1^* .$$

5. FUNCTIONS OF CONTROLS DEFINED BY ORDINARY DIFFERENTIAL EQUATIONS

Proofs. We shall use the notation of section 3 and we shall make, at first, the same assumptions as in Theorem 3.3.

Let, for any integrable function f from T to some euclidean space, T' (f) be the set of all the points t* in T such that $|f(t^*)|$ is finite and

$$\lim_{j \to \infty} \frac{1}{|M_j|} \int_{M_j} f(t)dt = f(t^*)$$

for all sequences $\{M_j\}_{j=1}^{\infty}$ of closed subsets of T that are regular at t*. It is well known [8, Th. (6.3), p. 118] that the set T' (f) has measure $|T|$.

Now let D^o be a bounded convex open set containing the range of $\bar{y}$, let $\psi_o = \psi^o_D$ be defined as in Assumption 3.2, let D_∞ be a dense denumerable subset of D^o and R_∞ a dense denumerable subset of $\bigcup_{t\in T} R^{\#}(t)$, and let $g(v, \cdot, \bar{\rho}(\ ))$ be the function $t \to g(v, t, \bar{\rho}(t))$. Then

$$T^* = \bigcap_{v\in D_\infty,\, r\in R_\infty} (T'(g(v,\cdot, r)) \cap T'(g(v,\cdot,\bar{\rho}(\cdot)))) \cap T'(\psi_o) \cap [t_o, t_1)$$

has measure $|T|$. We also verify that $T'(g(v,\cdot, r)) \supset T^*$ and $T'(g(v,\cdot,\bar{\rho}(\cdot))) \supset T^*$ for all $v\in D^o$ and $r\in R_\infty$. Indeed, let $t^*\in T^*$, $v\in D^o$, $r\in R_\infty$, and let $v_1, v_2, \ldots$ be a sequence in D_∞ converging to v. Then, by Assumption 3.2,

$$\limsup_{j\to\infty} \left| \frac{1}{|M_j|} \int_{M_j} g(v, t, r)dt - g(v_i, t^*, r) \right| \leq$$

$$\limsup_{j\to\infty} \frac{1}{|M_j|} \int_{M_j} |g(v, t, r) - g(v_i, t, r)|\, dt \leq$$

$$\limsup_{j\to\infty} |v - v_i| \frac{1}{|M_j|} \int_{M_j} \psi_o(t)dt = \psi_o(t^*) \left| v-v_i \right| \; (i= 1, 2, \ldots),$$

for every sequence $\{M_j\}$ that is regular at t^*. Since $g(\cdot, t^*, r)$ is continuous on D^o, we conclude that $T'(g(v,\cdot, r)) \supset T^*$. We similarly show that $T'(g(v,\cdot,\bar{\rho}(\cdot))) \supset T^*$ for all $v\in D^o$.

We next define sets $N_k(t, \alpha)$ and the corresponding collection $\mathscr{N}$. Let $m = n^2$, and let

$$\eta_{k-1,i} = 2^{-mi+m-k+1}, \quad N_k = \bigcup_{i=1}^{\infty} (\eta_{k,i}, \eta_{k-1,i}],$$

and

$$N_k(t) = (t + N_k) \cap T (t\in T;, k= 1, \ldots, m;\ i= 1, 2, \ldots).$$

We then define $N_k(t, \alpha)$ for $\alpha \geq 0$ as $N_k(t) \cap (t, t + \beta_k(t, \alpha)]$, where $\beta_k(t, \alpha)$ is nonnegative and such that $|N_k(t, \alpha)| = \mathrm{Min}(\alpha, |N_k(t)|)$. We observe that diameter $(N_k(t, \alpha)) \leq 2^m |N_k(t, \alpha)|$ for all t and α, and $|N_k(t, \alpha)| = \alpha$ for sufficiently small α and $t < t_1$.

We shall henceforth use the above definitions of T* and $\mathscr{N}$, as well as the definitions of $\mathscr{R}$* and R* of section 3. We shall also use the notation of Definition 2.1. Let $t^{ij} \in T^*$ and $\rho^{ij} \in \mathscr{R}^*(t^{ij})$ $(i, j = 1, \ldots, n)$, and let

$$\rho' = \rho'(t^{\square}, \rho^{\square}, \omega^{\square}) = [\rho^{ij}, N_{nj-n+i}(t^{ij}, \omega^{ij}) \ (i, j = 1, \ldots, n); \bar{\rho}]$$

for all $\omega^{\square} \in \Omega(t^{\square})$. Finally, let $B = B_o^*, \bar{b} = \bar{b}_o^*$, and for $\rho \in \mathscr{R}$ and $b \in B$, let the absolutely continuous function $y = y(\cdot\,; \rho, b)\colon T \to E_n$ be the solution of the system

$$\dot{y}(t) = g(y(t), t, \rho(t)) \quad \text{a.e. in } T,$$

$$y(t_o) = \phi_o(b).$$

It follows from Assumption 3.2 and from well known theorems that there exists a neighborhood $\tilde{B}$ of $\bar{b}$ in B such that the function y, as just defined, exists, is unique, has its range contained in D^o, and depends continuously on b, uniformly in ρ, for all $b \in \tilde{B}$ and all "measurable" ρ such that the set $\{t \in T \mid \rho(t) \neq \bar{\rho}(t)\}$ has a sufficiently small measure.

<u>Lemma 5.1.</u> Let $\rho' = \rho'(t^{\square}, \rho^{\square}, \omega^{\square})$. For all $t \in T$, $b \in \tilde{B}$, $t^{ij} \in T^*$, and $\rho^{ij} \in \mathscr{R}^*(t^{ij})$ $(i, j = 1, \ldots, n)$, the function $(\omega^{\square}, b) \to y(t; \rho', b)$ is continuous in some neighborhood Γ of $(0^{\square}, \bar{b})$ in $\Omega(t^{\square}) \times \tilde{B}$; and, for all $i, j = 1, \ldots, n$, the limit defining the right hand derivative of $y(t; \rho', b)$ with respect to ω^{ij} at 0 exists and is uniform in Γ, and this derivative is a continuous function of $(\omega^{\square}, b)$ in Γ. Similarly,

$$\lim_{\theta \to +0} \frac{1}{\theta} (y(t; \rho', b_1 + \theta(b_2 - b_1)) - y(t; \rho', b_1))$$

defining the right hand derivative

$$\partial y(t;\rho', b_1 + \theta(b_2 - b_1))/\partial\theta\Big|_{\theta=0}$$

exists, this limit is uniform in $\Gamma \times \Gamma$, and it is continuous in $\Gamma \times \Gamma$.

Finally, let $x(\rho, b) = y(t_1;\rho, b)$. Then

$$D_{\mathscr{N}_k} x(\bar{\rho}, \bar{b}; t^*, \rho^*) = Dx(\bar{\rho}, \bar{b}; t^*, \rho^*(t^*)) = Z(t^*)\,(g(\bar{y}(t^*), t^*, \rho^*(t^*)) - g(\bar{y}(t^*),$$

$$t^*, \bar{\rho}(t^*)))\ (k=1, 2, \ldots, n^2,\ t^* \in T^*, \rho^* \in \mathscr{R}^*(t^*)),$$

and

$$Dx(\bar{\rho}, \bar{b}; b) = Z(t_o)\,\frac{\partial\phi_o(\bar{b})}{\partial b}\,(b-\bar{b})\ (b \in B),$$

where the matrix function Z is the solution of the system

$$\dot{Z}(t) = -Z(t)\,g_v(\bar{y}(t), t, \bar{\rho}(t)) \quad \text{a.e. in } T,\ Z(t_1) = I \quad \text{(the unit matrix).}$$

Proof of Lemma 5.1. Let $t^\square$ and $\rho^\square$ be fixed. For fixed $\bar{i}$ and $\bar{j}$ in $\{1, 2, \ldots, n\}$, let $t^* = t^{\bar{i}\bar{j}}$, $\rho^* = \rho^{\bar{i}\bar{j}}$, and $M(\alpha) = N_{n\bar{j}-n+\bar{i}}(t^*, \alpha)$ for $\alpha \geq 0$.

We observe that, for every sequence $\alpha_1, \alpha_2, \ldots$ converging to $+0$, the sequence $\{\bar{M}_a\} = \{\bar{M}(\alpha_a)\}_{a=1}^{\infty}$ is regular at t^*, $|\bar{M}_a - M_a| = 0$ for all a, and $|M(\alpha)| = \alpha$ for sufficiently small α. It follows that, for all $v \in D^o$,

$$\lim_{\alpha \to +0} \frac{1}{\alpha} \int_{M(\alpha)} g(v, t, \rho(t))dt = g(v, t^*, \rho(t^*))$$

if $\rho = \bar{\rho}, \rho = \rho^*$, or $\rho(t) = r \in R_\infty$ on T.

We next consider $y(t;\rho', b)$ as a function of $(\omega^\square, b)$. The measure of the set $\{t\epsilon T \mid \rho'(t^\square, \rho^\square, \bar{\omega}^\square) \neq \rho'(t^\square, \rho^\square, \omega^\square)\}$ converges to 0 uniformly in $\Omega(t^\square)$ when

$$|\omega^\square - \bar{\omega}^\square| = \sum_{i,j=1}^{n} |\omega^{ij} - \bar{\omega}^{ij}| \to 0.$$

Furthermore, $|g(v, \cdot, r)|$ and $|g_v(v, \cdot, r)|$ are bounded by some integrable function ψ_1 on $D^o \times R$. We conclude, using standard arguments, that $y(t;\rho', b)$ is a uniformly continuous function of $(\omega^\square, b, t)$ and $y(t;\rho', b) \epsilon D^o$ in $\Gamma \times T$, where Γ is some neighborhood of $(0^\square, \bar{b})$ in $\Omega(t^\square) \times B$.

Now we fix b and sufficiently small $\omega^{ij}((i, j) \neq (\bar{i}, \bar{j}))$ as well as $\bar{i}, \bar{j}, t^\square$, and $\rho^\square$, and set $\tilde{\rho}(\alpha) = \rho'(t^\square, \rho^\square, \omega^\square)$ and $\tilde{y}(t;\alpha) = y(t;\tilde{\rho}(\alpha), b)$ for $\omega^{\bar{i}\bar{j}} = \alpha \geqslant 0$ and $t \epsilon T$. Then, for sufficiently small α, $\tilde{\rho}(t;\alpha) = \tilde{\rho}(t;0)$ for $t \epsilon T - M(\alpha)$, $\tilde{\rho}(t;\alpha) = \rho^*(t)$ and $\tilde{\rho}(t;0) = \bar{\rho}(t)$ for $t \epsilon M(\alpha)$, $\tilde{y}(t;\alpha) = \tilde{y}(t;0)$ for $t \leqslant t^*$, and, for $t > t^*$,

$$\Delta(t;\alpha) = \frac{1}{\alpha}(\tilde{y}(t;\alpha) - y(t;0)) =$$

$$\frac{1}{\alpha}\int_{t^*}^{t} (g(\tilde{y}(\theta;\alpha), \theta, \tilde{\rho}(\theta;\alpha)) - g(\tilde{y}(\theta;0), \theta, \tilde{\rho}(\theta;0)))\, d\theta;$$

hence

$$\text{(5.1.1)} \quad \Delta(t;\alpha) = \int_{[t^*, t] - M(\alpha)} g_v(\tilde{y}(\theta;0), \theta, \tilde{\rho}(\theta;0))\, \Delta(\theta;\alpha)\, d\theta$$

$$+ \frac{1}{\alpha}\int_{M(\alpha)} (g(\tilde{y}(\theta;\alpha), \theta, \rho^*(\theta)) - g(\tilde{y}(\theta;0), \theta, \bar{\rho}(\theta)))d\theta +$$

$$+ \int_{[t^*, t] - M(\alpha)} (g_v(\tilde{\tilde{y}}(\theta;\alpha), \theta, \tilde{\rho}(\theta;0)) - g_v(\tilde{y}(\theta;0), \theta, \tilde{\rho}(\theta;0)))\, \Delta(\theta;\alpha)d\theta,$$

where $\tilde{\tilde{y}}(\theta;\alpha)$ is, for each θ and α, intermediate between $\tilde{y}(\theta;\alpha)$ and $\tilde{y}(\theta;0)$. Since $\tilde{y}(\cdot\,;\alpha)$ converges uniformly to $\tilde{y}(\cdot\,;0)$ as $\alpha \to 0$, $|g_v(v,t,r)| \leqslant \psi_0(t)$ on $D^o \times T \times R$, $\rho^* \in R^*(t^*)$, and $T^* \subset T'(\psi_0)$, we can assert that

$$\lim_{\alpha \to +0} \frac{1}{\alpha} \int_{M(\alpha)} (g(\tilde{y}(\theta;\alpha), \theta, \rho^*(\theta)) - g(\tilde{y}(\theta;0), \theta, \bar{\rho}(\theta)))d\theta$$

$$= \lim_{\alpha \to +0} \frac{1}{\alpha} \int_{M(\alpha)} (g(\tilde{y}(t^*;0), \theta, \rho^*(\theta)) - g(\tilde{y}(t^*;0), \theta, \bar{\rho}(\theta)))d\theta$$

$$= g(\tilde{y}(t^*,0), t^*, \rho^*(t^*)) - g(\tilde{y}(t^*;0), t^*, \bar{\rho}(t^*)),$$

and that this limit is uniform in $\Gamma \times T$. Furthermore,

$$g_v(\tilde{\tilde{y}}(\theta;\alpha), \theta, \tilde{\rho}(\theta;0)) - g_v(\tilde{y}(\theta;0), \theta, \tilde{\rho}(\theta;0)) \mid$$

converges to 0 with α, for each fixed θ in T, uniformly in Γ, because $g_v(\cdot, t, r)$ is continuous, hence uniformly continuous, in some compact set containing D^o, for every t and r. Moreover, the uniform convergence, hence also the boundedness, of the second term on the right of (5.1.1) implies that $\Delta(\cdot\,;\cdot)$ is bounded. Since $|g_v(v,t,r)| \leq \psi_0(t)$ on $D^o \times T \times R$, it follows then that the last term on the right of (5.1.1) converges to 0 with α, uniformly in $\Gamma \times T$.

Let $\eta(t) = \lim_{\alpha \to +0} \Delta(t,\alpha)$ for $t \in T$. We can now conclude that η exists, is unique, that this limit is uniform in $\Gamma \times T$, and that

$$(5.1.2) \quad \eta(t) = \int_{t^*}^{t} g_v(\tilde{y}(\theta;0), \theta, \tilde{\rho}(\theta;0))\eta(\theta)d\theta$$

$$+ g(\tilde{y}(t^*,0), t^*, \rho^*(t^*)) - g(\tilde{y}(t^*,0), t^*, \bar{\rho}(t^*)) \quad \text{for } t > t^*.$$

Now we must investigate the dependence of η on $(\omega^{\square}, b)$. Let $(\omega_1^{\square}, b_1)$ and $(\omega_2^{\square}, b_2)$ be both in Γ and be such that $\omega_1^{\overline{ij}} = \omega_2^{\overline{ij}} = 0$, and let $y_1(\cdot), \rho_1(\cdot), \eta_1(\cdot)$ and $y_2(\cdot), \rho_2(\cdot), \eta_2(\cdot)$ represent the corresponding determinations of $\tilde{y}(\cdot\,;0)$, $\tilde{\rho}(\cdot\,;0)$, and $\eta(\cdot)$. Let also $M = \{t \in T \mid \rho_1(t) \neq \rho_2(t)\}$ and $\Delta(t) = |\eta_1(t) - \eta_2(t)|$. Then (5.1.2) yields

$$\begin{aligned}
\Delta(t) \leq & \int_M \psi_o(\theta)(\,|\eta_1(\theta)| + |\eta_2(\theta)|)d\theta + \int_{t^*}^{t} \psi_o(\theta)\Delta(\theta)\,d\theta \\
& + \int_T |g_v(y_1(\theta), \theta, \rho_1(\theta)) - g_v(y_2(\theta), \theta, \rho_1(\theta))| \cdot |\eta_2(\theta)|\,d\theta \\
& + |g(y_1(t^*), t^*, \rho^*(t^*)) - g(y_2(t^*), t^*, \rho^*(t^*))| \\
& + |g(y_1(t^*), t^*, \bar{\rho}(t^*)) - g(y_2(t^*), t^*, \bar{\rho}(t^*))| \qquad \text{for } t > t^*.
\end{aligned}$$

We can directly verify from (5.1.2) that η is uniformly bounded on $\Gamma \times T$. We can show, therefore, as in a previous argument, that the third integral in the last relation converges uniformly (on Γ) to 0 with $|\omega_1^{\square} - \omega_2^{\square}| + |b_1 - b_2|$. The first integral converges uniformly to 0 because $|M| \to 0$ uniformly, and the non-integrated terms converge to 0 uniformly with $|y_1(t^*) - y_2(t^*)|$. It follows that $\Delta(\cdot) \to 0$ uniformly on Γ as $|\omega_1^{\square} - \omega_2^{\square}| + |b_1 - b_2| \to 0$.

We can solve equation (5.1.2), specifically when $\omega^{\square} = 0^{\square}$ $b = \bar{b}$, and $r = \rho^*(t^*)$, and find that, for $k = n\bar{j} - n + \bar{i}$,

$$D_{\mathscr{N}_k} x(\bar{\rho}, \bar{b}; t^*, \rho^*) = \eta(t_1) = Z(t^*)(g(\bar{y}(t^*), t^*, r) - g(\bar{y}(t^*), t^*, \bar{\rho}(t^*))).$$

Thus $D_{\mathscr{N}_k} x$ is the same for all k and all $\rho^* \in \mathscr{R}^*(t^*)$ such that $\rho^*(t^*) = r$.

Similar arguments prove our assertions concerning $y(t; \rho', b_1 + \theta(b_2 - b_1))$ as a function of θ, and yield the representation of $Dx(\bar{\rho}, \bar{b}; b)$.

This completes the proof of the lemma.

5.2 Completion of the proof of Theorem 3.3. We shall now show that $(T^*, \mathscr{R}^*, \mathscr{N})$ define local variations for x in $\mathscr{R} \times B$ at $(\bar{\rho}, b)$. It is clear that, by construction, the collection $\mathscr{N}$ satisfies condition (2.1.1). Since the sets $N_k(t, \alpha)$ are unions of intervals and $\rho^{ij} \in \mathscr{R}$ $(i, j = 1, \ldots, n)$, the mapping ρ' belongs to $\mathscr{R}$.

It follows from Lemma 5.1 that the function $(\omega^{\square}, \theta^{\square}) \to \xi(\omega^{\square}, \theta^{\square})$ satisfies condition (2.1.2.2). Indeed, we have shown there that the right hand partial derivatives of ξ with respect to each ω^{ij} at $\omega^{ij} = 0$ and with respect to each θ^{ij} at $\theta^{ij} = 0$ exist, are continuous, and the limits defining them are uniform for $\omega^{\square}$ and $\theta^{\square}$ sufficiently close to $0^{\square}$. Finally, statement (2.1.2.3) follows directly from Lemma 5.1.

Thus $(T^*, \mathscr{R}^*, \mathscr{N})$ satisfy the conditions of Definition 2.1 and $Dx(\bar{\rho}, \bar{b}; t^*, r)$ and $Dx(\bar{\rho}, \bar{b}; b)$ have the representations described in Lemma 5.1. All the statements of Theorem 3.3, except statement (3.3.5), now follow directly from Theorem 2.2 after we set $z(t) = Z^T(t)\lambda$ on T. Furthermore, statement (2.2.2) implies (3.3.5), with $\bar{R}^*(t)$ replaced by $R^*(t)$. Since, however, $g(v, t, \cdot)$ is continuous on R for all v and t, we conclude that statement (3.3.5) is satisfied.

Finally, consider the special case when $R^{\#}(t) = R$ on T and $\mathscr{R}'$ contains all the constant mappings into R_∞. In that case, for each $r \in R_\infty$ and $t^* \in T^*$, the set $\mathscr{R}^*(t^*)$ contains the constant mapping from T to r, and $\bar{R}^*(t^*) = \bar{R}_\infty = R$.

This completes the proof of Theorem 3.3.

5.3 Proof of Theorem 3.5. The first part of the theorem, concerning the existence of $\bar{b}_o$ and $\bar{\sigma}$ as well as of the approximating sequences, follows directly from [1, Th. 3.1]. Next we observe that, for $S^{\#}(t) = \{s \in S \mid s(\bar{R}^{\#}(t)) = 1\}$ on T, S, $S^{\#}$, f, and $\mathscr{A}^{\#}$ satisfy the assumptions made in Theorem 3.3 about R, $R^{\#}$, g and $\mathscr{R}$, respectively. Furthermore, since $f(v, t, \tilde{s}) = g(v, t, \tilde{r})$ on $E_n \times T$ for every measure $\tilde{s} = s_{\tilde{r}}$ concentrated at the single point $\tilde{r}$, it follows that the set of σ in $\mathscr{S}^{\#}$ that are admissible (with respect to f and $S^{\#}$) at t* contains $\mathscr{R}^*(t^*)$. Finally, there exists a dense denumerable subset of S containing $\{s_r \mid r \in R_{\infty}\}$. We may now apply Theorem 3.3, with S, $S^{\#}$, f, and $\mathscr{S}^{\#}$ replacing R, $R^{\#}$, g, and $\mathscr{R}$, respectively, and derive directly the second part of Theorem 3.5.

REFERENCES

1. J. Warga, Functions of relaxed controls, J. SIAM Control 5 (1967).

2. ______ , Relaxed variational problems, J. Math. Anal. Appl. 4, (1962), pp. 111-128.

3. ______ , Necessary conditions for minimum in relaxed variational problems, J. Math. Anal. Appl. 4, (1962), pp. 129-145.

4. L.W. Neustadt, An abstract variational theory with applications to a broad class of optimization problems. I: General theory, S.I.A.M. J. Control 4, (1966).

5. ______ , An abstract variational theory with applications to a broad class of optimization problems II: Applications, S.I.A.M. J. Control 4, (1966).

6. E.J. McShane, Necessary conditions in generalized-curve problems of the calculus of variations, Duke Math, J. 7, (1940), pp. 1-27.

7. H. Halkin, On the necessary condition for optimal control of nonlinear systems, J. Analyse Math. 12, (1964), pp. 1-82.

8. S. Saks, Theory of the integral, Monografje Matematyczne, 2nd ed., 1937, reprinted by Hafner Publishing Co., New York.

9. N. Dunford and J. Schwarz, Linear Operators, Part I, Interscience, New York, 1964.

Northeastern University
Boston, Massachusetts

THE REDUCTION OF CERTAIN CONTROL PROBLEMS TO AN "ORDINARY DIFFERENTIAL" TYPE

By J. Warga
Professor, Northeastern University
Boston, Massachusetts

NASA Grant NGR 22-011-020
Supplement 1

THE REDUCTION OF CERTAIN CONTROL PROBLEMS

TO AN "ORDINARY DIFFERENTIAL" TYPE

by J. Warga*

The most commonly encountered problems of the mathematical control theory are of the "ordinary differential" type, that is, are defined by systems of ordinary differential equations involving control functions as well as by certain additional relations that must be satisfied by the control functions and the state variables. An interest has also been evidenced in certain more general problems in which, for example, ordinary differential equations are replaced by difference-differential, or, more generally, delay-differential equations, or in which absolutely continuous solutions of differential equations are replaced by piecewise absolutely continuous solutions with at most k jump discontinuities. The purpose of the present note is to show that many such non-standard problems can be easily transformed into an equivalent "ordinary differential" form to which all the, by now, classical results of the control theory are directly applicable.

Let U be an arbitrary set, E_n the euclidean n-space, $A \subset E_n$, $B \subset E_n \times E_n$, T the closed interval $[t_o, t_1]$ of the real axis, V a mapping from T into the class of nonempty subsets of U, and $f: E_n \times T \times U \to E_n$. An "ordinary differential" control problem consists in determining a function $u: T \to U$ and an absolutely continuous function $x = (x^1, x^2, .., x^n): T \to E_n$ such that $x^1(t_1)$ is minimum and

* This work was supported by N.A.S.A. grant NGR 22-011-020, Supplement 1.

$$(0.1)\quad \frac{dx(t)}{dt} = \dot{x}(t) = f(x(t),\ t,\ u(t)) \quad \text{a.e. in } T,$$

$$(0.2)\quad (x(t_o),\ x(t_1)) \in B,$$

$$(0.3)\quad x(t) \in A \ (t \in T),$$

and

$$(0.4)\quad u(t) \in V(t) \ (t \in T).$$

We shall now transform into an "ordinary differential" form a few of the more frequently encountered non-standard problems, namely (a) certain delay-differential control problems, (b) staging problems, and (c) problems involving variable times as well as constraints relating the values of x on a finite subset of T. For the sake of clarity, we consider only relatively simple versions of these problems.

1. Advance-delay-differential problems. Let τ_o, t_o, τ_1, and t_1 be fixed, $\tau_o < t_o < t_1 < \tau_1$, and let an absolutely continuous increasing function p be defined on $[\tau_o, \tau_1]$. Assume that $p(t) < t$, $\tau_o \le p(t_o)$, and $t_1 \le p(\tau_1)$. Let $p_o(t) \equiv t$ and $p_{i+1}(t) = p(p_i(t))$ $(i = 0, \pm 1, \pm 2, \ldots,\ p_i(t) \in [\tau_o, \tau_1])$. Let U, V, A, B, and T be defined as before, and let the functions x and u, into respectively E_n and U, be given on the intervals $[\tau_o, t_o)$ and $(t_1, \tau_1]$. Let $k(t)$ and $\ell(t)$ $(t \in [t_o, t_1])$ be integers defined by $p_{k(t)+1}(t) < \tau_o \le p_{k(t)}(t)$ and

$$p_{-\ell(t)}(t) \le \tau_1 < p_{-\ell(t)-1}(t).$$

(Such integers exist because $p(t)-t$ is negative and bounded away from 0. We also observe that k and ℓ are step-functions).

For any function y on $[\tau_o, \tau_1]$, let

$$\tilde{y}(t) = (y(p_{k(t)}(t)),\ y(p_{k(t)-1}(t)), \ldots,\ y(p_o(t)), \ldots, y(p_{-\ell(t)}(t)),$$

and let $f(\tilde{y},\tilde{v},t)$ be defined for every $t \in [t_o,t_1]$, every $\tilde{y} \in E_n \times\cdots\times E_n(k(t) + \ell(t) + 1$ times), and every $\tilde{v} \in U \times\cdots\times U(k(t) + \ell(t) + 1$ times).

We consider the problem of minimizing $x^1(t_1)$ subject to the functional equation

(1.1) $\dot{x}(t) = f(\tilde{x}(t), \tilde{u}(t), t)$ a.e. in $[t_o,t_1]$,

and subject to relations (0.2), (0.3), and (0.4). We also restrict x to be absolutely continuous on $[t_o,t_1]$. We shall show that this problem can be reduced to the "ordinary differential" type.

Let $\overline{k} = k(t_1)$, $\overline{\ell} = \ell(t_1) + 1$, $x_i(t) = x(p_i(t))$ and $u_i(t) = u(p_i(t))$ $(t \in (p(t_1),t_1],\ p_i(t) \in [\tau_o,\tau_1])$, $\hat{y} = (y_{\overline{k}},\ldots,y_{\underline{\overline{\ell}}})$ and $\hat{v} = (v_{\overline{k}},\ldots,v_{\underline{\overline{\ell}}})$ $(y_i \in E_n,\ v_i \in U)$, and let $\phi_i(\hat{y},\hat{v},t) = \dot{p}_i(t)\ f(\tilde{y},\tilde{v},t)$ $(t \in (p(t_1),t_1],\ i = 0,1,\ldots,\overline{k})$. Let also k' be a positive integer defined by $p_{k'}(t_1) \geq t_o > p_{k'+1}(t_1)$, and let $\overline{\tau} \in (p(t_1),t_1]$ be defined by $p_{k'}(\overline{\tau}) = t_o$. Let $y(t) = x_{k'}(t)$ on $[\overline{\tau},t_1]$ and $y(t) = x_{k'}(\overline{\tau})$ on $[p(t_1),\overline{\tau})$.

We now observe that our problem is equivalent to the "ordinary differential" problem of minimizing $x_o{}^1(t_1)$ subject to the differential equations

$$\dot{x}_i(t) = \phi_i(\hat{x}(t),\hat{u}(t),t) \text{ a.e. in } [p(t_1),t_1] \quad (i=0,1,\ldots,k'-1)$$

$$(1.2) \qquad \dot{y}(t) = \begin{cases} \dot{x}_{k'}(t) = \phi_{k'}(\hat{x}(t),\hat{u}(t),t) & \text{a.e. in } [\overline{\tau},t_1] \\ 0 & \text{a.e. in } [p(t_1),\overline{\tau}), \end{cases}$$

and subject to the restrictions

(1.3) $x_{i+1}(t_1) = x_i(p(t_1)) \quad (i = 0,1,\ldots,k'-1),$

(1.4) $(y(p(t_1)), x_o(t_1)) \varepsilon B, \quad y(t_1) = x_{k'}(t_1),$

(1.5) $x_i(t) \varepsilon A \quad (t \varepsilon [p(t_1),t_1], \; p_i(t) \varepsilon [t_o,t_1]),$

(1.6) $u_i(t) \varepsilon V(p_i(t)) \quad (t \varepsilon [p(t_1),t_1],$

$p_i(t) \varepsilon [t_o,t_1])$

2. Staging problems. We consider the "ordinary differential" control problem but, instead of restricting the curve x to be absolutely continuous on $[t_o,t_1]$, we require that x be piecewise absolutely continuous, with at most k jump discontinuities on the interval $[t_o,t_1]$. We also require that $x(\tau_i+)(=\lim_{\substack{t>\tau_i \\ t\to\tau_i}} x(t))$ and $x(\tau_i-)(=\lim_{\substack{t<\tau_i \\ t\to\tau_i}} x(t))$ satisfy relations of the form $(x(\tau_i+), x(\tau_i-)) \varepsilon B_i \subset E_n \times E_n$ at the i-th point of discontinuity τ_i in $[t_o,t_1]$. The values of $\tau_i(i=1,\ldots,k)$ can either be preassigned, or can be freely chosen. In the latter case, however, we assume that $V(t) = U$ for all t. A typical example of such a problem is the guidance of a rocket that can jettison certain "stages" when they are no longer needed.

We let

$$\tau_o = t_o, \quad \tau_{k+1} = t_1,$$
$$x_i(\theta) = x(\tau_i + \theta(\tau_{i+1}-\tau_o)) \quad (0<\theta<1, \; i=0,1,\ldots,k),$$
$$x_i(0) = \lim_{\theta\to0} x_i(\theta), \quad x_i(1) = \lim_{\theta\to1} x_i(\theta),$$
$$u_i(\theta) = u(\tau_i + \theta(\tau_{i+1} - \tau_o)) \quad (0 < \theta < 1, \; i=0,1,\ldots,k).$$

If the $\tau_i(i=0,1,\ldots, k+1)$ are preassigned, our problem becomes one of minimizing $x_k^1(1)$ by a choice of an absolutely continuous function $x = (x_o, x_1,\ldots,x_k)$ on $[0,1]$ satisfying the relations

$$(2.1)\quad \frac{dx_i(\theta)}{d\theta} = (\tau_{i+1} - \tau_i)\, f(x_i(\theta),$$
$$\tau_i + \theta(\tau_{i+1} - \tau_i),\ u_i(\theta))$$
$$\text{a.e. in } [0,1] \quad (i=0,1,\dots,k)$$

$$(2.2)\quad (x_o(0),\ x_k(1)) \in B$$

$$(2.3)\quad (x_i(1),\ x_{i+1}(0)) \in B_i \quad (i=0,1,\dots,k-1)$$

$$(2.4)\quad x_i(\theta) \in A \quad (0 \le \theta \le 1)$$

$$(2.5)\quad u_i(\theta) \in V(\tau_i + \theta(\tau_{i+1} - \tau_o)) \quad (0<\theta<1,\ i=0,1,\dots,k)$$

If the τ_i can be freely chosen in some preassigned interval [a,b], we may treat them as constant functions on [0,1], and adjoin the following relations:

$$\frac{d\tau_i}{d\theta} = 0 \quad \text{a.e. in } [0,1] \quad (1 = 0,\dots,k)$$

$$a \le \tau_o(0) \le \dots \le \tau_{k+1}(0) \le b.$$

3. Variable times, and constraints on a finite subset of $[t_o,t_1]$.

We now modify the standard "ordinary differential" problem as follows:

Let $\tau_i (i=1,..,k)$ be either preassigned or free to choose subject to the relations $t_o \le \tau_o \le \dots \le \tau_k \le t_1$. In addition to relations (0.1), (0.2), (0.3), and (0.4), we require that the absolutely continuous function x satisfy the relation

$$(x(t_o),\ x(\tau_1),\dots,\ x(\tau_k),\ x(t_1)) \in B^{\#},$$

where $B^{\#}$ is a given set.

We may proceed as we did for the "staging" problem, except that the relations (2.2) and (2.3) are replaced by

$$(3.1) \quad (x_o(0), x_1(0), \ldots, x_k(0), x_k(1)) \varepsilon B^{\#}$$

and

$$(3.2) \quad x_i(1) = x_{i+1}(0) \quad (i=0,1,\ldots,k-1)$$

Northeastern University
Boston, Massachusetts

A PRIMER ON THE PRIMER

By P. M. Lion
Assistant Professor of Aerospace and Mechanical Sciences
Princeton University
Princeton, New Jersey

* * * * * * *

A PRIMER ON THE PRIMER

By P. M. Lion

Assistant Professor of Aerospace and Mechanical Sciences
Princeton University
Princeton, New Jersey

SUMMARY

* * * * * * *

The necessary conditions for an impulsive trajectory to be optimum can be stated in terms of Lawden's primer vector. Recently, the primer vector has also been shown to have significance for non-optimal trajectories, indicating how these trajectories can be improved. This paper presents a simplified derivation of both results from a single viewpoint. In addition, a computational scheme for determining optimum n-impulse trajectories is suggested.

Introduction

The term "primer vector" was introduced by Lawden (Ref. 1) to denote the three adjoint variables associated with the velocity vector on an optimal trajectory. Lawden derived a necessary condition for the optimality of impulsive trajectories in terms of the magnitude of this vector. (Optimum in this memo is defined as minimum characteristic velocity.)

Recently (Ref. 2), the definition of the primer vector has been extended to non-optimal impulsive trajectories. It can then be shown that the primer gives a clear indication of how the original, or reference, trajectory can be improved; i.e., how the reference trajectory can be altered so as to decrease the total characteristic velocity while still satisfying the boundary conditions. The two main results of (Ref. 2) are

(1) the criterion for an additional impulse. Using this test indicates whether or not the reference trajectory can be improved by an additional midcourse impulse.

(2) the transversality condition. Using this test one can determine how the interior (midcourse) impulses of the reference trajectory should

be moved in both position and time so as to decrease the characteristic velocity. In addition, it is possible to determine whether initial and/or final coasts will improve the trajectory.

In this paper the results of (Ref. 1) and (Ref. 2) are rederived from a single viewpoint. In addition, a computational scheme for determining optimum n-impulse trajectories is suggested.

Formulation and Notation

The equations of motion are

$$\ddot{x} = \nabla\Phi \ (x, t) \tag{1}$$

where Φ is the gravitational potential. For impulsive trajectories, it is assumed that the velocity vector $v = (v_1, v_2, v_3)$ can be altered discontinuously; however, the position vector $x = (x_1, x_2, x_3)$ must be continuous. The criterion of optimality is the sum of the magnitudes of the velocity increments,

$$J = \sum_k |\Delta v_k|$$

The optimization problem may then be stated as follows: given an initial state (v_o, x_o) and a final state (v_f, x_f), find the trajectory which connects these states in a given travel time, t_f, such that J is minimized.

Assume that some trajectory Γ has been found which satisfies the boundary conditions and consider small perturbations about Γ. Let (x, v) and (x', v') denote the state vectors on Γ and on the perturbed trajectory Γ' respectively. Define

$$\begin{aligned} \delta x(t) &= x'(t) - x(t) \\ \delta v(t) &= v'(t) - v(t) \end{aligned} \tag{2}$$

If Γ and Γ' are sufficiently close to justify a linear analysis, then $(\delta v, \delta x)$ are, to first order, the solutions of the following variational equations of (1):

$$\begin{pmatrix} \delta\dot{x} \\ \delta\dot{v} \end{pmatrix} = \begin{pmatrix} O & I \\ G & O \end{pmatrix} \begin{pmatrix} \delta x \\ \delta v \end{pmatrix} \tag{3}$$

where I is the (3 x 3) identity matrix and G is the "gravity gradient" matrix. The elements of G are given by

$$g_{ij} = \frac{\partial^2 \Phi}{\partial x_i \, \partial x_j}$$

In second order form (3) can be written

$$\delta \ddot{x} = G \, \delta x \tag{4}$$

The (6 x 6) transition matrix $\Omega(t, \tau)$ for this system can be partitioned into four (3 x 3) matrices as follows:

$$\Omega(t, \tau) = \begin{pmatrix} \Omega_{11}(t, \tau) & \Omega_{12}(t, \tau) \\ \\ \Omega_{21}(t, \tau) & \Omega_{22}(t, \tau) \end{pmatrix} \tag{5}$$

The adjoint system to (3) is

$$\begin{pmatrix} \dot{\mu} \\ \\ \dot{\lambda} \end{pmatrix} = \begin{pmatrix} O & -G \\ \\ -I & O \end{pmatrix} \begin{pmatrix} \mu \\ \\ \lambda \end{pmatrix}$$

where μ and λ are 3 - vectors. In second order form this becomes

$$\ddot{\lambda} = G\lambda \tag{6}$$

identical to (4). Hence the transition matrix for $(\lambda, \dot{\lambda})$ will be identical to (5):

$$\begin{pmatrix} \lambda(t) \\ \\ \dot{\lambda}(t) \end{pmatrix} = \Omega(t, \tau) \begin{pmatrix} \lambda(\tau) \\ \\ \dot{\lambda}(\tau) \end{pmatrix} \tag{7}$$

It can be shown by differentiation that the identity

$$\lambda \cdot \delta v - \dot{\lambda} \cdot \delta x = \text{constant} \tag{8}$$

holds everywhere on Γ. This equation is the basis for most of the analysis which follows.

Consider Γ, a two impulse trajectory (or two-impulse segment of a multi-impulse trajectory), with impulses Δv_o at t_o and Δv_f at t_f. The primer vector λ is defined as the solution to system (6) which satisfies the following boundary conditions:

$$\lambda(t_o) = \lambda_o = \frac{\Delta v_o}{|\Delta v_o|}$$

$$\lambda(t_f) = \lambda_f = \frac{\Delta v_f}{|\Delta v_f|}$$

That is, at the endpoints of Γ, λ has unit magnitude and is aligned with the velocity increment. A solution satisfying these boundary conditions can be found if $\Omega_{12}(t_f, t_o)$ is non-singular; the initial value of $\dot{\lambda}$ is given by

$$\dot{\lambda}_o = \Omega_{12}^{-1}(t_f, t_o)(\lambda_f - \Omega_{11}(t_f, t_o)\lambda_o)$$

The above definition of the primer vector is extended easily to multi-impulse trajectories. In such cases, the right hand side of (6) is different on the different segments. At each impulse point, t_k, λ is again defined as a unit vector in the direction of the impulse

$$\lambda(t_k) = \frac{\Delta v_k}{|\Delta v_k|}$$

The solutions from different arcs are, therefore, patched together so that λ is continuous over the entire trajectory. The primer rate $\dot{\lambda}$ will, in general, be discontinuous at impulse points. (On optimal trajectories, however, it will be shown that $\dot{\lambda}$ is continuous and is orthogonal to λ at interior impulses.)

Criterion for Additional Impulse

Consider the two impulse trajectory Γ (shown schematically in Figure 2) which goes from x_o to x_f in the prescribed transit time. Γ may be a complete trajectory or a two impulse segment of a multi-impulse trajectory. By Lambert's theorem (Ref. 3) there are no other two impulse trajectories (in the neighborhood of Γ) which satisfy these boundary conditions. There is, however, a four parameter family of three impulse trajectories which do satisfy these conditions.

Assume that Γ passes through the point x_m at $t = t_m$. The four parameters used to describe a neighboring three impulse trajectory Γ' will be the time of the midcourse impulse, t_m, and the position relative to Γ at this time $\delta x_m = x'(t_m) - x(t_m)$.

Γ' is constructed as follows: The impulse $(\Delta v_o + \delta v_o)$ is applied at t_o so that Γ' will pass through $x_m + \delta x_m$ at t_m. Γ' must be continuous in position, but not in velocity. A small midcourse impulse $(\delta v_m^+ - \delta v_m^-)$ is required to null the position displacement at $t_f (\delta x_f = 0)$. Finally, the impulse $(\Delta v_f + \delta v_f)$ is applied at t_f so that the final velocity is matched.

The costs on Γ and Γ', dropping the higher order terms, are as follows:

on Γ : $J = |\Delta v_o| + |\Delta v_f|$

on Γ' : $J' = |\Delta v_o + \delta v_o| + |\delta v_m^+ - \delta v_m^-| + |\Delta v_f + \delta v_f|$

The difference in cost, to first order, is

$$\delta J = \frac{\Delta v_o}{|\Delta v_o|} \cdot \delta v_o + |\delta v_m^+ - \delta v_m^-| - \frac{\Delta v_f}{|\Delta v_f|} \cdot \delta v_f$$

From the definition of the primer vector,

$$\delta J = \lambda_o \cdot \delta v_o + |\delta v_m^+ - \delta v_m^-| - \lambda_f \cdot \delta v_f$$

Using (8) this becomes

$$\delta J = -\lambda_m \cdot (\delta v_m^+ - \delta v_m^-) + |\delta v_m^+ - \delta v_m^-|$$

This expression is homogeneous in $(\delta v_m^+ - \delta v_m^-)$.

Denoting the magnitude of the midcourse impulse by c,

$$\delta J = c\,(1 - \lambda_m \cdot \eta)$$

where η is a unit vector in the direction of $(\delta v_m^+ - \delta v_m^-)$.

If δJ can be made negative, then Γ' represents an improvement in cost over Γ. This can occur if, and only if, $p(t_m) > 1$. That is, if $p(t_m) > 1$ it is possible, by varying the direction of δx_m, to find a Γ' such that the required midcourse impulse $(\delta v_m^+ - \delta v_m^-)$ will point in the direction of λ_m. Clearly, for this choice δJ is negative.

Therefore we have the following results:

(a) If Γ is an optimum trajectory (or a segment of a multi-impulse optimum trajectory), it is necessary that $p(t) \leq 1$ for all t in the interval (t_o, t_f).

(b) If $p(t) > 1$ for any t in the interval (t_o, t_f), then there exists a neighboring trajectory with an additional impulse which lowers the cost. To first order, the greatest improvement in cost can be realized by applying the midcourse impulse at the time the primer magnitude reaches its maximum, and in the direction of the primer vector.

If the time history of the primer vector is as shown in Figure 2, then the necessary condition for optimality is satisfied. Figure 3 shows an example of a case where an additional impulse improves the trajectory.

These conclusions are the result of a linear analysis, and, therefore, effects are additive. If, then, the primer magnitude exceeds unity at more than one point, the reference trajectory can be improved (at least to first order) by adding impulses at both points.

Finally, if the reference trajectory is two impulse, then there is no other (neighboring) two impulse trajectory which satisfies the boundary conditions. If $p(t) < 1$ for all $t_o < t < t_f$, then all neighboring trajectories with more impulses actually <u>increase</u> the cost. In this case, then, we have sufficient conditions for the reference trajectory to be optimal. Figure 2 represents such a case.

Transversality

In this section, an expression is developed which gives the differential cost between two neighboring trajectories. This expression is the analog of the transversality condition of the calculus of variations. In this case, however, there is wider applicability since neither trajectory need be an optimum. This is in distinction to the finite thrust case, and is a result of the fact that "cost" is incurred only at discrete points.

Consider the two neighboring trajectories shown in Figure 4. Both trajectories are initially on orbit C_o at t_o. On Γ, impulses are applied at t_o and t_f. Γ', on the other hand, remains on C_o until t_1 $(= t_o + dt_o)$ and then impulses are applied at t_1 and t_f. Both trajectories have terminal state corresponding to orbit C_f at t_f.

The symbol $d(\cdot)$ will be used to indicate noncontemporaneous variations, that is

$$dx = x'(t + dt) - x(t)$$

To first order the relationship between dx and δx is

$$dx = \delta x + \dot{x}\, dt\,.$$

On Γ, the cost is

$$J = |v_o^+ - v_o^-| + |v_f^+ - v_f^-|$$

and on Γ'

$$J' = |v_1^+ - v_1^-| + |v_f^{\pm}(v_f^- + \delta v_f)|$$

The differential cost is given by

$$\delta J = \lambda_o \cdot (dv_o^+ - dv_o^-) - \lambda_f \cdot \delta v_f^- \tag{9}$$

where $dv_o^+ = v_1^+(t_o + dt_o) - v_o^+(t_o)$

$$= \delta v_o^+ + \dot{v}_o^+ dt_o$$

and $dv_o^- = \dot{v}_o^- dt_o$

Since $\dot{v}$ is continuous (it depends only on position and time)

$$dv_o^+ - dv_o^- = \delta v_o^+$$

Substituting in (9) and using (8)

$$\delta J = \dot{\lambda}_o \cdot \delta x_o$$

or, since $\delta x_o = dx_o - \dot{x}_o^+ dt_o$

$$\delta J = \dot{\lambda}_o \cdot dx_o - (\dot{\lambda}_o \cdot \dot{x}_o^+)\, dt_o \tag{10}$$

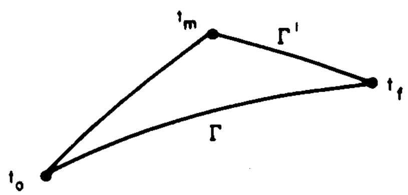


Figure 1

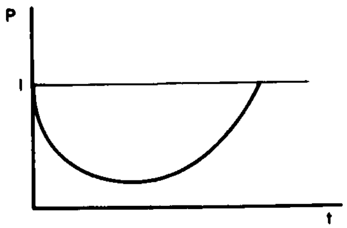


Figure 2

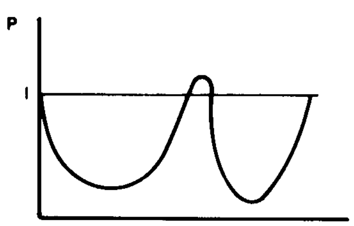


Figure 3

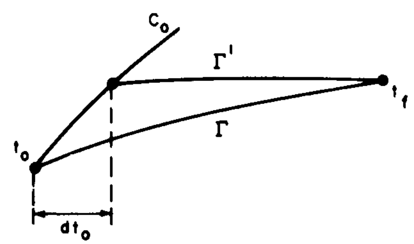


Figure 4

This "transversality condition" represents the difference in cost between two neighboring trajectories whose initial points differ in position by dx_o and in time by dt_o (and whose final position and time are identical). Note that no considerations of optimality have been required.

If the final point differs by (dt_f, dx_f), the proper expression is

$$\delta J = -\dot{\lambda}_f \cdot dx_f + (\dot{\lambda}_f \cdot \dot{x}_f^+)\, dt_f \tag{11}$$

Equation (10) can be put in more familiar form by adding $\lambda_o \cdot (\dot{v}_o^+ - \dot{v}_o^-) = 0$ and noting that $dv_o = \dot{v}_o^- \, dt_o$. Equation (10) then becomes

$$\delta J = -\lambda_o \cdot dv_o + \dot{\lambda}_o \cdot dx_o + H\, dt_o \tag{12}$$

which is exactly the same form as the usual transversality equation for finite thrust (optimal) trajectories. Similarly (11) becomes

$$\delta J = \lambda_f \cdot dv_f - \dot{\lambda}_f \cdot dx_f - H\, dt_f \tag{13}$$

Equations (10) and (11) are the fundamental form, however, since the differentials are independent.

Final and Initial Coasts

To test the desirability of an initial coast, the reference trajectory is compared with a neighboring trajectory which has been allowed to coast $(dt_o > 0)$ in the initial orbit (Figure 4).

Using the transversality condition (10), the difference in cost to first order is

$$\delta J = \dot{\lambda}_o \cdot dx_o - (\dot{\lambda}_o \cdot \dot{x}_o^+)\, dt_o$$

Substituting $dx_o = \dot{x}_o^- \, dt_o$, (the superscript minus refers to increments along the initial orbit), this becomes

$$\delta J = \dot{\lambda}_o \cdot (\dot{x}_o^- - \dot{x}_o^+)\, dt_o$$

Since $dt_o > 0$, δJ will be negative if

$$\dot{\lambda}_o \cdot (\dot{x}_o^- - \dot{x}_o^+) < 0$$

From the definition of the primer λ_o is parallel to Δv_o. Therefore, the last equation implies

$$-(\dot{\lambda}_o \cdot \lambda_o) < 0$$

or

$$\left.\frac{dp}{dt}\right|_{t=0} > 0$$

In other words, if the primer magnitude exceeds unity immediately after the initial impulse, an initial coast would lower the cost. Figure 5 is an example of this.

Similarly, it can be shown that if

$$\left.\frac{dp}{dt}\right|_{t=t_f} < 0$$

then a final coast improves the cost. The trajectories being compared in this case are shown in Figure 6. For this case, it must be remembered that, to meet the time constraint, $dt_f < 0$. The primer in Figure 7 is an example.

Figure 8 represents a trajectory which is so far from optimal that almost anything (additional impulse, initial coast, final coast) will improve it.

Circular Coplanar Orbits

In the special case of transfers between circular coplanar orbits the above conditions have a simple geometric interpretation.

It is more convenient here to shift to polar coordinates. Equation (13) for a final coast becomes

$$\delta J = \lambda_\theta \, d\theta_f^+ - H \, dt_f^+ < 0$$

and for an initial coast, Equation (12) becomes

$$\delta J = -\lambda_\theta \, d\theta_o^- + H \, dt_o^- < 0$$

where $\lambda_\theta = \dot{\lambda}_1 x_2 - \dot{\lambda}_2 x_1 - \lambda_1 \dot{x}_2 + \lambda_2 \dot{x}_1$.

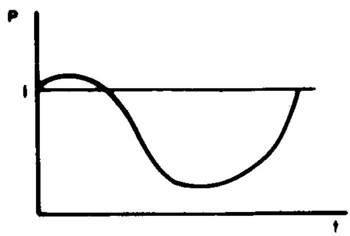


Figure 5

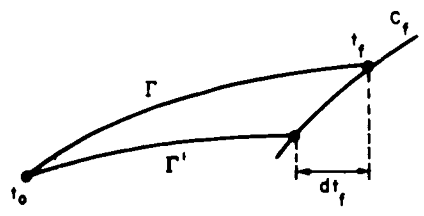


Figure 6

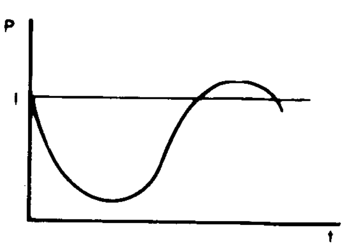


Figure 7

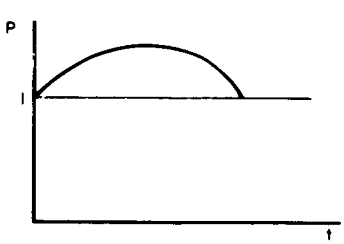


Figure 8

Let ω_o be the angular rate of the initial orbit and ω_f the angular rate of the final orbit. Then the above conditions can be written

$$\begin{aligned} &\text{Final coast} \quad (dt_f < 0): \; H/\lambda_\theta > \omega_f \\ &\text{Initial coast} \quad (dt_o > 0): \; H/\lambda_\theta < \omega_o \end{aligned} \tag{14}$$

Now consider Figure 9. Here contours of constant J are plotted against the prescribed central angle (φ) and trip time (T). These contours are closed about the minimum J, or Hohmann, trajectory. The value of J increases going outward from this point.

Along the contours, since J is constant,

$$\delta J = \lambda_\theta \, d\varphi - H \, dt = 0$$

(Note: $\lambda_\theta = \frac{\partial J}{\partial \theta_f}$, $H = -\frac{\partial J}{\partial t_f}$)

The slope of the contours is thus given by

$$m = \frac{d\varphi}{dT} = H/\lambda_\theta$$

To interpret equations (14) geometrically consider Figure 10. The original two-impulse contours are shown as broken lines. Straight lines with slope ω_o and ω_f have been drawn tangent to these contours. The solid lines represent contours of constant J when initial and final coasts are considered. In region A, where $m < \omega_o$, initial coasts represent an improvement. In region C, where $m > \omega_f$, final coasts are an improvement.

For example, if a transfer corresponding to point P_1 is required, it is cheaper to coast initially through angle $\Delta\varphi$ (which takes time ΔT) and then perform the transfer corresponding to P_2.

Also note that the set of points on the φ - T plane which can be reached for the minimum (Hohmann) cost is a wedge (cross-hatched on the figure). Any point in this wedge can be reached by a unique combination of initial and final coasts plus the 180^o transfer. Other points may need either an initial or final coast, but not both.

Moving Interior Impulses

Consider the two three impulse segments shown in Figure 11. The differential cost between these two trajectories can be derived from Equations (10) and (11) and is given by

$$\delta J = [\dot{\lambda}_m^+ \cdot dx_m - (\dot{\lambda}_m^+ \cdot x_m^+)dt_m] - [\dot{\lambda}_m^- \cdot dx_m - (\dot{\lambda}_m^- \cdot x_m^-)dt_m]$$

$$= (\dot{\lambda}_m^+ - \dot{\lambda}_m) \cdot dx_m + (H^+ - H^-)dt_m \tag{15}$$

where the equation

$$(H^+ - H^-) = -(\lambda_m^+ \cdot \dot{x}^+ - \lambda_m^- \cdot \dot{x}^-) \tag{16}$$

has been used. dx_m and dt_m in Equation (15) can be chosen independently.

The following conclusions can then be drawn

(a) If the reference trajectory is an optimum, then it is necessary that

$$\dot{\lambda}^+ = \dot{\lambda}^- \quad \text{and} \quad H^+ = H^-$$

Since H is constant on any segment, it is therefore constant along the entire trajectory.

If these two functions are continuous, then

$$\dot{\lambda}_m \cdot (\dot{x}_m^+ - \dot{x}_m^-) = 0$$

Since λ_m is aligned with the velocity impulse this last equation becomes

$$\dot{\lambda}_m \cdot \lambda_m = 0$$

or $\frac{dp}{dt} = 0$ at impulse points.

(b) If Γ is not an optimum, then a neighboring trajectory with lower cost can be found by choosing

$$\begin{aligned} dx_m &= -\epsilon\, [\dot{\lambda}_m^+ - \dot{\lambda}_m^-] \\ dt_m &= -\epsilon\, [H^+ - H^-] \end{aligned} \tag{17}$$

If ϵ is "sufficiently small," then the three impulse trajectory which passes through $x_m + \epsilon\, dx_m$ at $t_m + \epsilon\, dt_m$ will represent an improvement over Γ. Therefore, Equation (15) tells us how interior impulses should be moved in position and time so as to reduce characteristic velocity.

A Gradient Scheme for Optimum Multi-Impulse Trajectories

In this section, a technique is suggested for determining optimum n-impulse trajectories (where n is open) starting with the two impulse, or "Lambert," solution (or any other nominal trajectory).

This technique is based upon two results: (1) the criterion for an additional impulse, which tells when an additional impulse should be added, and (2) the transversality equation as developed in Equation (15) which tells how interior impulses should be moved.

A necessary part of this technique is a subroutine which solves Lambert's problem; i.e., given both position and velocity at two different times, find the trajectory which connects them. Also, it is desirable to have a subroutine which computes the transition matrix $\Omega(t, \tau)$ (Equation (5)). For the inverse square field, the formulation of the Lambert problems by Pines (Ref. 4) and the transition matrix routine by Goodyear (Ref. 5) represent elegant answers to these needs. The formulation of both is done in the "universal variables" and thus is valid without modification for all conic sections.

The iteration procedure is as follows: given the position and velocity at two terminals plus transit time, determine from the Lambert subroutine the two impulse trajectory which connects them. Imposing the appropriate boundary conditions on the primer, determine the time history of $p(t)$. If the primer magnitude appears as in Figure 2, then the two impulse trajectory is at least a local optimum. If the primer magnitude rises above 1.0, then a third impulse must be added. Let t_m indicate the time when the primer reaches a maximum. Using the boundary conditions $\delta x(t_o) = 0$, $\delta x(t_f) = 0$, it can be shown that for any trajectory Γ' passing through $x(t_m) + \delta x_m$,

$$(\delta v_m^{+} - \delta v_m^{-}) = A\, \delta x_m$$

where $$A = \Omega_{22}(t_m, t_f)\, \Omega_{12}^{-1}(t_m, t_f) - \Omega_{22}(t_m, t_o)\, \Omega_{12}^{-1}(t_m, t_o)$$

For maximum improvement, $(\delta v_m^+ - \delta v_m^-)$ should be parallel to λ_m. Therefore, assuming A is nonsingular, choose

$$\delta x_m = \epsilon A^{-1} \lambda_m$$

where ϵ is a small constant to insure that the range of linearity is not violated.

Two Lambert problems are then solved.

(1) Connecting (x_o, t_o) with $(x_m + \delta x_m, t_m)$

(2) Connecting $(x_m + \delta x_m, t_m)$ with (x_f, t_f)

If ϵ is small enough, then this three impulse trajectory will represent an improvement. In all probability the three impulse trajectory is not an optimum. For instance, the plot of primer magnitude vs time may look as shown in Figure 12. In this event, after the $\dot{\lambda}$ and H before and after the impulses are calculated, the following corrections are made as given by Equation (17):

$$dx_m = -\epsilon (\dot{\lambda}_m^+ - \dot{\lambda}_m^-)$$

$$dt_m = -\epsilon (H^+ - H^-)$$

Again two Lambert problems are calculated and the process repeated until

$$| H^+ - H^- | < \eta$$

$$| \lambda_m^+ - \lambda_m^- | < \zeta$$

where η and ζ are preassigned tolerances.

If at any point in the iteration, the primer magnitude becomes greater than 1.0 at some point, then an additional impulse is added. From that point on, three Lambert problems must be solved at each iteration. The interior impulses are then moved in the same manner as above. In principle, there is no limit to the number of impulses which can be handled by this formulation.

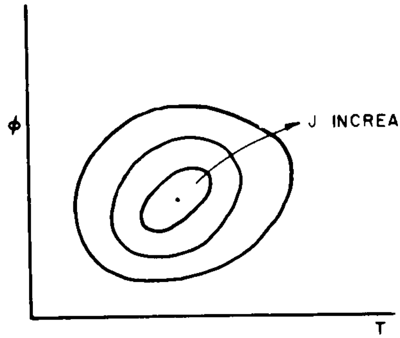


Figure 9

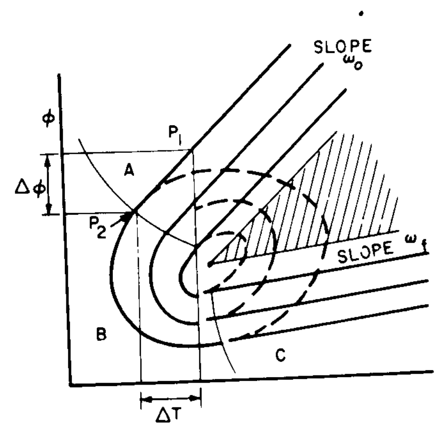


Figure 10

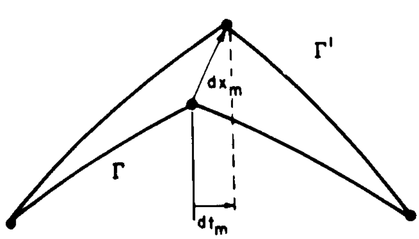


Figure 11

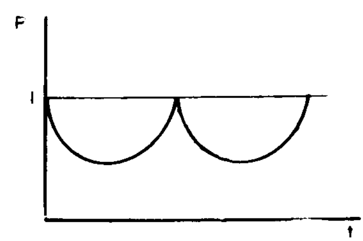


Figure 12

This technique, which is actually a gradient computation, exhibits the properties of first order techniques in general: guaranteed improvement on each iteration but convergence slowing as the minimum is approached.

The method is presently being programmed by the ASMAR group at Princeton.

References

1. Lawden, D. F.: Optimal Trajectories for Space Navigation, Butterworths, London, 1963.

2. Lion, P. M. and Handelsman, M.: "The Primer Vector on Fixed Time Impulsive Trajectories," AIAA paper 67-54.

3. Battin, R. H.: Astronautical Guidance, McGraw-Hill, 1964.

4. Pines, S.: "A Uniform Solution of Lambert's Problem," MSC Internal Note No. 65-FM-106.

5. Goodyear, W. H.: "A General Method for the Computation of Cartesian Coordinates and Partial Derivatives of the Two Body Problem," NASA CR-522, 1965.

GROUP THEORETICAL ASPECTS

OF THE PERTURBATION OF KEPLERIAN MOTION

By D. C. Lewis and Pinchas Mendelson
Control Research Associates
Baltimore, Maryland

NASA Contract NAS 12-93

GROUP THEORETICAL ASPECTS OF THE PERTURBATION OF KEPLERIAN MOTION

by

Daniel C. Lewis
Professor of Applied Mathematics
The Johns Hopkins University
Baltimore, Maryland

and

Pinchas Mendelson
Associate Professor of Mathematics
Brooklyn Polytechnic Institute
Brooklyn, New York

SUMMARY

The following principle is exploited to obtain five linearly independent solutions of the variational equations for Keplerian motion. The principle: If a system of differential equations is invariant under a continuous and differentiable group of transformations, it is possible in general, by differentiations only, to write down a number of linearly independent solutions of the variational equations equal to the number of independent parameters of the group. In the Keplerian case there is, however, a removable singularity occurring when the motion is circular.

A sixth solution of the variational equations is given by differentiation with respect to the eccentricity e, or rather with respect to $\cos^{-1}e$ in the elliptic case and with respect to $\cosh^{-1}e$ in the hyperbolic case. A more complicated function of e can be used in the parabolic case, a parabola being thought of as the limit of a family of non parabolic conics.

Numerous formulas and identities are written out explicitly for manipulations in the elliptic case. A variation of the Lagrange method for integrating non-homogeneous linear differential equations, especially adapted for systems of second order, is developed and applied to the elliptic case of Keplerian motion.

Finally it is shown how the Keplerian differential equations are invariant under a group of transformations which, in general, change the eccentricity provided that one allows the independent variable to undergo a differential transformation.

INTRODUCTION

The purpose of this paper is to collect in as explicit a form as possible the formulas necessary for handling the perturbation of Keplerian motion in rectangular coordinates. The emphasis is on elliptic motion although brief mention is also made of both hyperbolic and parabolic motion.

The proposed method is one of successive approximations based on the repeated integration (by quadratures) of a series of non-homogeneous variational equations. It is more explicit and automatic than the Brouwer method (Cf. Brouwer and Clemence, Methods of Celestial Mechanics, Academic Press 1961, pages 398 to 414), but in other respects is much the same thing. The Brouwer method itself contains little originality, since, for example, his method of integrating the non-homogeneous variational equations is due in principle to Lagrange. The merits of the present paper are therefore more in the realm of explicitness and exposition than of originality, although it contains a number of formulas and theorems that we have not found elsewhere.

The method of **Lagrange** is based on a prior knowledge of a complete set of solutions of the homogeneous variational equations. The method of finding these is based on a group theoretical principle which has contributed to the title of the paper and which is explained as follows:

Consider a system of differential equations of the form

$$x^* = f(t, x) \ , \tag{1}$$

where x and f are n-vectors, t is the independent variable, f is of class C', and x^* represents either the first or the second derivative of x with respect to t (or, for that matter, the result of any linear homogeneous differential operator with constant coefficients acting on x). The homogeneous variational system based on a given solution $x(t)$ is, by definition, the system.

$$\xi^* = A(t)\xi \ ,$$

where $A(t)$ is an $n \times n$-matrix, namely the jacobian matrix of the components of f with respect to the components of x, with x replaced by $x(t)$. Now, if $x(t)$ can be imbedded in a one parameter family of solutions, say $x(t, p)$, in such wise that $x(t, p_0) = x(t)$, it is both well known and obvious that a solution of the variational equations is given by $\xi = (\partial x/\partial p)$ with p set equal to p_0. Group theory is of value in finding a way to imbed the given solution in a family of solutions depending on one or more parameters. Suppose that (1) is invariant under a group of transformations whose equations are of the form

$$x' = h(t, x, p), \quad t' = P(t, x, p) .$$

Assume that p_0 is the parameter value corresponding to the identity, so that $h(t, x, p_0) \equiv x$ and $P(t, x, p_0) \equiv t$. Then any fixed solution of (1), gives rise to a solution $x(t, p)$ such that

$$x(P(t, x(t), p), p) \equiv h(t, x(t), p)$$

a formula which will lead to an explicit expression for $x(t', p)$ provided that the equation $t' = P(t, x(t), p)$ can be solved for t in terms of t' and p. In any case, it is clear from the above identity (by setting $p = p_0$) that $x(t, p_0) = x(t)$. Several very simple examples of the utility of these considerations are given in the next Section in connection with the variationa equations of Keplerian motion.

I. SOLUTIONS OF THE VARIATIONAL EQUATIONS

The equations for Keplerian motion, with proper choice of the units, may be written in the form

$$\frac{d^2x}{dt^2} = -\frac{x}{[x^2+y^2+z^2]^{3/2}}$$

$$\frac{d^2y}{dt^2} = -\frac{y}{[x^2+y^2+z^2]^{3/2}}$$

$$\frac{d^2z}{dt^2} = -\frac{z}{[x^2+y^2+z^2]^{3/2}}$$

The equations of variation based on a given solution of these Keplerian equations, say, $x = \phi(t)$, $y = \psi(t)$, $z = \omega(t)$, take the form

$$\frac{d^2\xi}{dt^2} = \frac{(2\phi^2-\psi^2-\omega^2)\xi + 3\phi\psi\eta + 3\phi\omega\zeta}{[\phi^2+\psi^2+\omega^2]^{5/2}}$$

$$\frac{d^2\eta}{dt^2} = \frac{3\psi\phi\xi + (2\psi^2-\phi^2-\omega^2)\eta + 3\psi\omega\zeta}{[\phi^2+\psi^2+\omega^2]^{5/2}}$$

$$\frac{d^2\zeta}{dt^2} = \frac{3\omega\phi\xi + 3\omega\psi\eta + (2\omega^2-\phi^2-\psi^2)\zeta}{[\phi^2+\psi^2+\omega^2]^{5/2}}$$

We propose, using the group-theoretical principle explained above to find a complete set of six solutions of the variational equations based on any solution of a one parameter family of solutions of the Keplerian equations, say

$$x = \phi(\lambda, t), \quad y = \psi(\lambda, t), \quad z = \omega(\lambda, t) ,$$

where the parameter λ is assumed to be independent of the group parameters.

The <u>six</u> solutions are exhibited as the six columns in the following 3×6-matrix

$$W^* = \begin{bmatrix} \phi - (3/2)t\,\dot{\phi} & \dot{\phi} & 0 & \omega & -\psi & \phi_\lambda \\ \psi - (3/2)t\,\dot{\psi} & \dot{\psi} & -\omega & 0 & \phi & \psi_\lambda \\ \omega - (3/2)t\,\dot{\omega} & \dot{\omega} & \psi & -\phi & 0 & \omega_\lambda \end{bmatrix}$$

where the dots denote differentiation with respect to t.

The solution in the first column comes from what we shall call the scale group. Namely the Keplerian equations are invariant under the group of transformations $x' = p^{-1}x$, $y' = p^{-1}y$, $z' = p^{-1}z$, $t' = p^{-3/2}t$ in which the identity occurs when $p = 1$. Thus a solution of the Keplerian equations is $x = p\,\phi(\lambda, p^{-3/2}t)$, $y = p\,\psi(\lambda, p^{-3/2}t)$, $z = p\,\omega(\lambda, p^{-3/2}t)$. Differentiating with respect to p and then setting $p = 1$, we get the elements in the first column of the matrix W^*.

The second column comes from the autonomous group, the equations for which are $x' = x$, $y' = y$, $z' = z$, $t' = t+p$. Thus a solution of the Keplerian equations is $x = \phi(\lambda, t+p)$, $y = \psi(\lambda, t+p)$, $z = \omega(\lambda, t+p)$. Differentiating with respect to p and then setting $p = 0$ (which corresponds to the identity of the present group), we get the elements in the second column of W^*.

The third, fourth, and fifth columns come from the rotation groups. For instance the Keplerian equations are invariant under the group $x' = x$, $y' = y\cos p - z\sin p$, $z' = y\sin p + z\cos p$, $t' = t$, that is the group of rotations about the x-axis. Thus a solution of the Keplerian equations is

$x = \phi$, $y = \psi \cos p - \omega \sin p$, $z = \psi \sin p + \omega \cos p$. Differentiating with respect to p and setting $p = 0$, we get the third column of W^*. The fourth and fifth columns are obtained in a similar way by consideration of rotation groups about the y- and z-axes.

Finally the sixth column is certainly a solution of the variational equations because it consists merely of the derivatives of ϕ, ψ, and ω with respect to the parameter λ.

It is known that Keplerian motion always takes place in a plane. Hence there is no essential loss of generality in choosing the coordinate system in such a way that the given Keplerian motion takes place in the xy-plane. This amounts to taking $\omega \equiv 0$; thus greatly simplifying the matrix W^* by annihilating five of its elements (in addition to the three elements which are already zero). Furthermore, it is obvious, both from the columns of W^* and from the variational equations themselves, that with $\omega \equiv 0$, the variational system splits into two systems one involving only ξ and η and the other involving only ζ. The second of these systems is relatively trivial, and hence from this point on we confine most of our attention to the first of these systems, namely the system

$$
\frac{d^2\xi}{dt^2} = \frac{(2\phi^2-\psi^2)\xi + 3\phi\psi\eta}{(\phi^2+\psi^2)^{5/2}}
\qquad (2)
$$

$$
\frac{d^2\eta}{dt^2} = \frac{3\phi\psi\xi + (2\psi^2-\phi^2)\eta}{(\phi^2+\psi^2)^{5/2}}
$$

where ϕ and ψ satisfy the two dimensional Keplerian equations,

$$
(3) \qquad \ddot{\phi} = \frac{-\phi}{(\phi^2+\psi^2)^{3/2}}, \qquad \ddot{\psi} = \frac{-\psi}{(\phi^2+\psi^2)^{3/2}}
$$

The most general elliptic case can be obtained (by the operation of a scale transformation, a rotation, and a translation of t) from the following one-parameter family of solutions:

$$
(4) \qquad \begin{aligned} \phi &= -\cos\lambda + \cos\theta \\ \psi &= \sin\lambda \sin\theta \end{aligned}
$$

where θ is to be regarded as a function of t through the Kepler equation

$$
(5) \qquad t = \theta - \cos\lambda \sin\theta.
$$

An elementary calculation shows that (3) is satisfied by (4) and (5). The quantity θ is known as the eccentric anomaly. The eccentricity of the orbit is $\cos\lambda$, at least, if $0 < \lambda \leq \pi/2$.

Similarly the hyperbolic case can be handled by taking

$$\phi = -\cosh\lambda + \cosh\theta$$

$$\psi = \sinh\lambda \sinh\theta$$

where θ is defined as a function of t by the equation

$$t = \cosh\lambda \sinh\theta - \theta.$$

The most general parabolic case can be treated by taking

$$\phi = \frac{1}{2}(1 - \theta^2)$$

$$\psi = \theta$$

where θ is defined as a function of t by means of the equation

$$t = \frac{1}{2}(\theta - \frac{1}{3}\theta^3) .$$

The parabolic case differs from the other two, because all parabolas are obtainable from this one standard parabola by the operation of the three groups. There is no parameter λ, independent of the group parameters, enabling us to write the last column of the matrix W^*. This situation can be rectified by finding a family of conics consisting of ellipses and/or hyperbolas which, as the parameter λ of the family tends to a certain value, say, 0, converges to our standard parabola. Such a family of ellipses converging to our parabola as $\lambda \to 0$ is the following:

$$\phi = -\cot\lambda \csc\lambda + \csc^2\lambda \cos(\theta \sin\lambda)$$

$$\psi = \csc\lambda \sin(\theta \sin\lambda)$$

$$t = \theta \csc^2\lambda - \cot\lambda \csc^2\lambda \sin(\theta \sin\lambda) .$$

It is not within the scope of this paper to discuss any further either the parabolic or hyperbolic orbits.

II. MISCELLANEOUS FORMULAS FOR ELLIPTIC MOTION

We restrict most of our attention to the elliptic case. Since we are also restricting attention to the planar case, the equations of variation will have only four linearly independent solutions. They are exhibited as the four columns in the following 2×4-matrix,

$$W = \begin{bmatrix} \phi - (3/2)t\,\dot{\phi} & \dot{\phi} & -\psi & \phi_\lambda \\ \psi - (3/2)t\,\dot{\psi} & \dot{\psi} & \phi & \psi_\lambda \end{bmatrix} \tag{6}$$

This matrix is, of course, derivable from the matrix W^* by omitting the last row and the third and fourth columns. In writing down the explicit expressions for these solutions, it must be remembered that θ is regarded as a function of both t and λ. We readily find from (5) that

$$\frac{\partial\theta}{\partial t} = \frac{1}{1 - \cos\lambda\cos\theta} \tag{7}$$

and

$$\frac{\partial\theta}{\partial\lambda} = \frac{-\sin\lambda\sin\theta}{1 - \cos\lambda\cos\theta} \tag{8}$$

Thus, we find from (4) that

$$\dot{\phi}(t) = -\sin\theta(\partial\theta/\partial t) = \frac{-\sin\theta}{1 - \cos\lambda\cos\theta} \tag{9}$$

$$\dot{\psi}(t) = +\sin\lambda\cos\theta\,(\partial\theta/\partial t) = \frac{\sin\lambda\cos\theta}{1 - \cos\lambda\cos\theta} \tag{10}$$

$$\phi_\lambda(t) = \sin\lambda - \sin\theta\,(\partial\theta/\partial\lambda) = \sin\lambda + \frac{\sin\lambda\sin^2\theta}{1 - \cos\lambda\cos\theta} \tag{11}$$

$$\begin{aligned} \psi_\lambda(t) &= \cos\lambda\sin\theta + \sin\lambda\cos\theta\,(\partial\theta/\partial\lambda) \\ &= \cos\lambda\sin\theta - \frac{\sin^2\lambda\cos\theta\sin\theta}{1 - \cos\lambda\cos\theta} \end{aligned} \tag{12}$$

These formulas are sufficient to enable us to write down at once all the elements of W. It will also be necessary for us to have $\dot{W}$, the derivative of W with respect to t. For this purpose we record the following formulas obtained by differentiating formulas (9)-(10) with respect to t.

$$\ddot{\phi}(t) = \frac{\cos\lambda - \cos\theta}{(1 - \cos\lambda\cos\theta)^3} \tag{13}$$

(14) $$\ddot{\psi}(t) = \frac{-\sin\lambda\sin\theta}{(1-\cos\lambda\cos\theta)^3}$$

Before differentiating ϕ_λ and ψ_λ it is convenient first to observe that they may also be expressed in the forms

(15) $$\phi_\lambda(t) = \sin\lambda - \dot{\phi}(t)\ \psi(t)$$

(16) $$\psi_\lambda(t) = \cos\lambda\sin\theta - \psi(t)\ \dot{\psi}(t)$$

Hence

(17) $$\dot{\phi}_\lambda(t) = -\ddot{\phi}(t)\ \psi(t) - \dot{\phi}(t)\ \dot{\psi}(t)\ .$$

To find $\dot{\psi}_\lambda(t)$ first note from (4) that

(18) $$\phi^2 + \psi^2 = (\ 1 - \cos\lambda\cos\theta)^2$$

Differentiating this, we get

$$\phi\dot{\phi} + \psi\dot{\psi} = (1-\cos\lambda\ \cos\theta)(\cos\lambda\sin\theta)(\partial\theta/\partial t) = \cos\lambda\sin\theta$$

So (16) may be written

(19) $$\psi_\lambda(t) = \phi\dot{\phi} + \psi\dot{\psi} - \psi\dot{\psi} = \phi\dot{\phi}$$

Hence

(20) $$\dot{\psi}_\lambda(t) = \dot{\phi}^2 + \phi\ddot{\phi}$$

The $\ddot{\phi}$ and $\ddot{\psi}$ may always be eliminated with the help of (3).

Some other formulas worth noting and easily derivable from the preceding are the following

(21) $$\phi(t)\phi_\lambda(t) + \psi(t)\psi_\lambda(t) = \phi(t)\sin\lambda$$

(22) $$\phi(t)\dot{\psi}(t) - \dot{\phi}(t)\psi(t) = \sin\lambda$$

(23) $$\frac{2}{(\phi^2+\psi^2)^{1/2}} - (\dot{\phi}^2 + \dot{\psi}^2) = 1\ .$$

If we write the variational system (2) as a system of four linear differential equations of the first order

$$\frac{d\xi}{dt} = 0\,\xi + 0\,\eta + 1\,\dot{\xi} + 0\,\dot{\eta}$$

$$\frac{d\eta}{dt} = 0\,\xi + 0\,\eta + 0\,\dot{\xi} + 1\,\dot{\eta}$$

$$\frac{d\dot{\xi}}{dt} = A\,\xi + B\,\eta + 0\,\dot{\xi} + 0\,\dot{\eta}$$

$$\frac{d\dot{\eta}}{dt} = C\,\xi + D\,\eta + 0\,\dot{\xi} + 0\,\dot{\eta}$$

where $A = \dfrac{2\phi^2 - \psi^2}{(\phi^2 + \psi^2)^{5/2}}$, $B = + C = \dfrac{3\phi\psi}{(\phi^2 + \psi^2)^{5/2}}$ and $D = \dfrac{(2\psi^2 - \phi^2)}{(\phi^2 + \psi^2)^{5/2}}$,

it is seen at once that the trace of the matrix of coefficients on the right is zero. Hence, by a known theorem on linear differential equations, we know that the determinant of the matrix $\begin{pmatrix} W \\ \dot{W} \end{pmatrix}$ is a constant. It can therefore be evaluated in a simple manner by evaluating it at $t = 0$. From the above formulas it is easy to write down the following matrix.

$$\text{(24)} \qquad \begin{bmatrix} W(0) \\ \dot{W}(0) \end{bmatrix} = \begin{bmatrix} 1 - \cos\lambda & 0 & 0 & \sin\lambda \\ 0 & \dfrac{\sin\lambda}{1 - \cos\lambda} & 1 - \cos\lambda & 0 \\ 0 & -\dfrac{1}{(1-\cos\lambda)^2} & -\dfrac{\sin\lambda}{1 - \cos\lambda} & 0 \\ -\dfrac{1}{2}\left(\dfrac{\sin\lambda}{1-\cos\lambda}\right) & 0 & 0 & \dfrac{-1}{1 - \cos\lambda} \end{bmatrix}$$

Evidently

$$\det\begin{pmatrix} W(0) \\ \dot{W}(0) \end{pmatrix} = \begin{vmatrix} 1 - \cos\lambda & \sin\lambda \\ -\dfrac{1}{2}\left(\dfrac{\sin\lambda}{1-\cos\lambda}\right) & \dfrac{-1}{1-\cos\lambda} \end{vmatrix} \cdot \begin{vmatrix} \dfrac{\sin\lambda}{1-\cos\lambda} & 1 - \cos\lambda \\ \dfrac{-1}{(1-\cos\lambda)^2} & \dfrac{-\sin\lambda}{(1-\cos\lambda)} \end{vmatrix}$$

$$= [-(1/2)(1-\cos\lambda)][-\cos\lambda(1-\cos\lambda)^{-1}] = + 2^{-1}\cos\lambda.$$

This shows that our four solutions are not linearly independent if $\lambda = \pi/2$. In fact this becomes obvious anyway; for, with $\lambda = \pi/2$, $\theta = t$, $\phi(t) = \cos t$, $\psi(t) = \sin t$, $\dot{\phi}(t) = -\psi(t)$, $\dot{\psi}(t) = \phi(t)$, so that the second and third columns of W are identical. The case $\lambda = \pi/2$ is the case of zero eccentricity, corresponding to circular motion. The singularity occurring at $\lambda = \pi/2$ is, however, easily removable, in accordance with the following theorem.

<u>Theorem 1.</u> The variational equations admit the solution

$$\xi = \frac{\dot{\phi} + \psi}{\cos \lambda}\ , \qquad \eta = \frac{\dot{\psi} - \phi}{\cos \lambda}$$

and this solution has a removable singularity at $\lambda = \pi/2$.

<u>Proof.</u> The first statement is an obvious consequence of the fact that the (ξ,η) of the theorem is a linear combination of the second and third columns of the matrix W of (6).

As for the second statement, we find by an elementary calculation based on (4), (9), and (10) that

$$\xi = \frac{\dot{\phi} + \psi}{\cos \lambda} = \frac{1}{\cos \lambda}\left[\frac{-\sin \theta}{1 - \cos\lambda \cos\theta} + \sin\lambda \sin\theta\right]$$

$$= \frac{-\sin\theta + \sin\lambda \sin\theta - \sin\lambda \cos\lambda \sin\theta \cos\theta}{(1 - \cos\lambda \cos\theta)\cos \lambda}$$

$$= \left(\frac{-\sin \theta}{1 - \cos\lambda \cos\theta}\right)\left(\frac{1 - \sin\lambda}{\cos \lambda}\right) - \frac{\sin\lambda \sin\theta \cos\theta}{(1 - \cos\lambda \cos\theta)}\ .$$

$$\eta = \frac{\dot{\psi} - \phi}{\cos \lambda} = \frac{1}{\cos \lambda}\left[\frac{\sin\lambda \cos\theta}{1 - \cos\lambda \cos\theta} - (-\cos\lambda + \cos\theta)\right]$$

$$= \frac{\sin\lambda \cos\theta - \cos\theta + \cos\lambda + \cos\lambda \cos^2\theta - \cos^2\lambda \cos\theta}{(1 - \cos\lambda \cos\theta)\cos\lambda}$$

$$= -\left(\frac{\cos \theta}{1 - \cos\lambda \cos\theta}\right)\left(\frac{1 - \sin\lambda}{\cos \lambda}\right) + \left(\frac{1 + \cos^2\theta - \cos\lambda \cos\theta}{1 - \cos\lambda \cos\theta}\right)\ .$$

Now $\frac{1-\sin\lambda}{\cos\lambda} = \frac{\cos\lambda}{1+\sin\lambda} \to 0$ as $\lambda \to \pi/2$. We also have $\theta \to t$. Hence in the limit as $\lambda \to \pi/2$, we have

$$\xi = -\sin t \cos t = -\frac{1}{2}\sin 2t$$

$$\eta = 1 + \cos^2 t = \frac{3}{2} + \frac{1}{2}\cos 2t .$$

It is readily verified that these equations satisfy the variational equations when $\phi = \cos t$ and $\psi = \sin t$. They afford a solution which may be used to replace either the second or third column of W to preserve a complete set of linearly independent solutions in the case $\lambda = \pi/2$.

The inverse of the matrix in (24) is important for future purposes. It was calculated in the usual routine way, and we record the result here.

$$(25)\quad \begin{bmatrix} W(0) \\ \dot{W}(0) \end{bmatrix}^{-1} = \begin{bmatrix} \frac{2}{(1-\cos\lambda)^2} & 0 & 0 & \frac{2\sin\lambda}{1-\cos\lambda} \\ 0 & \tan\lambda & \frac{(1-\cos\lambda)^2}{\cos\lambda} & 0 \\ 0 & \frac{-1}{(1-\cos\lambda)\cos\lambda} & -\tan\lambda & 0 \\ \frac{-\sin\lambda}{(1-\cos\lambda)^2} & 0 & 0 & -2 \end{bmatrix}$$

III. APPLICATION OF NON-HOMOGENEOUS LINEAR EQUATIONS TO THE PERTURBATION OF ELLIPTIC MOTION

Perturbation of Keplerian elliptic motion generally requires the solution of a non-linear system, which, with suitable choice of coordinates, may be put in the form of the following three equations.

$$(26)\qquad \frac{d^2\xi}{dt^2} = \frac{(2\phi^2 - \psi^2)\xi + 3\phi\psi\eta}{[\phi^2 + \psi^2]^{5/2}} + f$$

$$(27) \qquad \frac{d^2\eta}{dt^2} = \frac{3\phi\psi\xi + (2\psi^2 - \phi^2)\eta}{[\phi^2 + \psi^2]^{5/2}} + g$$

$$(28) \qquad \frac{d^2\zeta}{dt} = \frac{-\zeta}{[\phi^2 + \psi^2]^{3/2}} + h$$

in which the non-linearities in ξ, η, and ζ occur only in f, g, and h and are small compared to the linear terms. Solutions are required to satisfy given initial conditions or more complicated boundary conditions. A useful method is that of successive approximations in which an approximate solution is inserted for ξ, η, ζ in f, g, h, which accordingly are temporarily regarded as known functions of t. We then are apt to get a better approximation by integrating the resulting non-homogeneous linear system, and then the process is repeated. The process, when infinitely repeated, converges to an exact solution under suitable conditions. In this paper, however, we are not concerned with convergence questions, but rather with efficient methods for solving the non-homogeneous linear system consisting of (26), (27), and (28) when f, g, and h are regarded as known functions of t.

Notice that (28) can be solved independently of (26) and (27). Indeed it admits the general solution

$$(29) \quad \zeta = a\phi(t) + b\psi(t) + \csc\lambda \int_0^t [\psi(t)\phi(s) - \phi(t)\psi(s)]h(s)ds ,$$

where a and b are constants of integration. The reader may verify this statement a posteriori with the help of (3) and (22).

We wish to get a similar result for (26) and (27), which are not decoupled. In the next section we develop a general theory of non-homogeneous linear systems consisting of n second order differential equations. When $n = 1$, this theory leads to the trivial result (29). When $n = 2$, it yields the not so trivial analogous result for the system consisting of (26) and (27).

IV. A MODIFICATION OF LAGRANGE'S THEORY OF VARIATION OF PARAMETERS

We wish to develop here a modification of the formula of Lagrange for solving a non-homogeneous system of linear differential equations in terms of a complete set of solutions of the corresponding homogeneous system. This formula is commonly referred to as the "Lagrange variation of parameters" formula. We wish to present a version which can be applied directly to a system of n equations of the second order in n unknowns without the necessity of rewriting the system as one of $2n$ equations of the first order in $2n$ unknowns. The system of interest in our immediate applications is of the form

$$\frac{d^2x}{dt^2} = A(t)x + f(t) \tag{30}$$

where x is an n-vector whose components are the n unknown functions, $A(t)$ is an $n \times n$-matrix whose elements are known continuous functions of the independent variable t, and f is an n-vector whose components are known continuous functions of t. It became evident however that the more general equation (35) below could be treated equally well. The $n \times n$-matrix $B(t)$ will be assumed to be of class C'.

Theorem 2. Suppose that $\Theta(t, s)$ is an $n \times n$-matrix whose elements are functions of class C'' of the two variables t and s. Suppose furthermore that

$$\ddot{\Theta}(t, s) = A(t)\Theta(t, s) + B(t)\dot{\Theta}(t, s) , \tag{31}$$

$$\Theta(s, s) = 0 \tag{32}$$

$$\dot{\Theta}(s, s) = I, \text{ the } n \times n\text{-identity matrix,} \tag{33}$$

and let

$$x(t) = \int_{t_0}^{t} \Theta(t, s)f(s)ds . \tag{34}$$

Then

$$\ddot{x}(t) = A(t)x(t) + B(t)\dot{x}(t) + f(t) \tag{35}$$

$$x(t_0) = 0 \tag{36}$$

$$\dot{x}(t_0) = 0 , \tag{37}$$

where the dot on Θ represents differentiation with respect to its first argument and the dot on $x(t)$ denotes differentiation with respect to t.

Proof. Evidently (36) follows at once from (34). Differentiating (34) we have

$$\dot{x}(t) = \int_{t_0}^{t} \dot{\Theta}(t, s) f(s)ds + \Theta(t, t)f(t).$$

Because of (32) this reduces to

$$\dot{x}(t) = \int_{t_0}^{t} \dot{\Theta}(t, s)f(s)ds \; . \tag{38}$$

and then (37) is seen to hold. Now differentiating (38), we get

$$\ddot{x}(t) = \int_{t_0}^{t} \ddot{\Theta}(t, s)f(s)ds + \dot{\Theta}(t, t)f(t) \; .$$

Hence from (31) and (33), we get

$$\ddot{x}(t) = \int_{t_0}^{t} [A(t)\Theta(t, s) + B(t)\dot{\Theta}(t, s)]f(s)ds + f(t) \; .$$

This can also be written

$$\ddot{x}(t) = A(t) \int_{t_0}^{t} \Theta(t, s)f(s)ds + B(t) \int_{t_0}^{t} \dot{\Theta}(t, s)f(s)ds + f(t).$$

Hence from (34) and (38) we see that (35) must hold.

The existence and uniqueness of the matrix $\Theta(t, s)$, with the properties described in the hypothesis of the above theorem, are self evident from the existence, uniquenoss, and continuity theorems for linear differential equations. It may be calculated from any $n \times 2n$-matrix solution X of the equation

$$\ddot{X}(t) = A(t)X(t) + B(t)\dot{X}(t) \; , \tag{39}$$

provided that

$$\det \begin{pmatrix} X \\ \dot{X} \end{pmatrix} \neq 0. \tag{40}$$

Since the columns of $\Theta(t, s)$ are solutions of $\ddot{x} = A(t)x + B(t)\dot{x}$, they must be linear combinations of the $2n$ columns of $X(t)$, with coefficients which are functions of s. This amounts to saying that there exists an $n \times 2n$-matrix $Y(s)$, such that

$$\Theta(t, s) = X(t)Y'(s) \qquad (Y' = \text{transpose of } Y) \; . \tag{41}$$

$Y(s)$ is now determined by (32) and (33), which may also be written

$$X(t)Y'(t) = 0 \tag{42}$$

$$\dot{X}(t)Y'(t) = I \; , \tag{43}$$

which last two equations, because of (40), are just sufficient for the determination of Y.

Theorem 3. The matrix $Y(t)$ described above satisfies the system

$$\ddot{Y}(t) = (A'(t) - \dot{B}'(t))Y(t) - B'(t)\dot{Y}(t) \ . \tag{44}$$

Proof. Differentiating (42), we obtain

$$\dot{X}Y' + X\dot{Y}' = 0 \tag{45}$$

Hence from (43), we have

$$\text{(i)} \quad \dot{X}Y' = I \qquad \text{and} \quad \text{(ii)} \qquad X\dot{Y}' = -I \ . \tag{46}$$

Differentiating (46i) and using (39), we have

$$[AX + B\dot{X}]Y' + \dot{X}\dot{Y}' = 0 \ . \tag{47}$$

Since $XY' = 0$ and $\dot{X}Y' = I$, we find from (47) that

$$\dot{X}\dot{Y}' = -B \ . \tag{48}$$

Differentiating this last relation and again using (39), we have

$$[AX + B\dot{X}]\dot{Y}' + \dot{X}\ddot{Y}' = -\dot{B} \tag{49}$$

But, since by (46) $X\dot{Y}' = -I$ and by (48) $\dot{X}\dot{Y}' = -B$, we find that (49) becomes $-A - B^2 + \dot{X}\ddot{Y}' = -\dot{B}$ so that

$$\dot{X}\ddot{Y}' = A + B^2 - \dot{B} \tag{50}$$

Multiplying (46i) on the right by $A - \dot{B}$ we have

$$\dot{X}Y'(A - \dot{B}) = A - \dot{B} \ , \tag{51}$$

while from (48) we have

$$\dot{X}\dot{Y}'B = -B^2 \ . \tag{52}$$

Subtracting (51) from (50) and adding (52) we get

$$\dot{X}[\ddot{Y}' - Y'(A - \dot{B}) + \dot{Y}'B] = 0 \tag{53}$$

Differentiating (46ii), we have $\dot{X}\dot{Y}' + X\ddot{Y}' = 0$, so that it follows from (48) that

$$X\ddot{Y}' = B \ . \tag{54}$$

We multiply (42) on the right by $(A - \dot{B})$ thus obtaining

$$XY'(A - \dot{B}) = 0. \tag{55}$$

Using, once again, (46ii), multiplying on the right by B, we see that

$$X\dot{Y}'B = - B\ . \tag{56}$$

Adding (54) and (56) and subtracting (55), we obtain

$$X[\ddot{Y}' - Y'(A - \dot{B}) + \dot{Y}'B] = 0\ . \tag{57}$$

It follows from (40), (53), and (57) that

$$\ddot{Y}' - Y'(A - \dot{B}) + \dot{Y}'B = 0\ .$$

Hence, taking the transpose, $\ddot{Y} - (A' - \dot{B}')Y + B'\dot{Y} = 0$, which is obviously equivalent to (44), as we wished to prove.

The system (35) is said to be self-adjoint if

$$A'(t) - \dot{B}' = A(t) \quad \text{and} \quad B'(t) = - B(t)\ . \tag{58}$$

In the self adjoint case $Y(t)$ and $X(t)$ satisfy the same differential system, namely (39).

<u>Theorem 4.</u> $Y(t)$ satisfies the initial conditions

$$\begin{aligned} &X(0)Y'(0) = 0 && X(0)\dot{Y}'(0) = - I \\ &\dot{X}(0)Y'(0) = I && \dot{X}(0)\dot{Y}'(0) = - B(0)\ , \end{aligned} \tag{59}$$

and, since (40) holds, $Y(t)$ is uniquely determined by (59) and (44).

<u>Proof.</u> We get (59) from (42), (46), and (48), by setting $t = 0$. Because of (40) with $t = 0$, the initial values of Y and $\dot{Y}$ are uniquely determined by (59), and Theorem 4 follows from existence and uniqueness theorems for differential equations.

<u>Theorem 5.</u> Let $L(t) = \begin{pmatrix} X(t) \\ \dot{X}(t) \end{pmatrix}$ and $J = \begin{pmatrix} 0 & -I \\ I & -B_0 \end{pmatrix}$ where $B_0 = B(0)$. Then, if the system is self adjoint, $Y(t)$ may be found from the formula

$$Y(t) = X(t)\, L_0^{-1}\, J'\, (L_0^{-1})'\ , \qquad L_0 = L(0). \tag{60}$$

and

(61) $$\Theta(t, s) = X(t)\, L_0^{-1}\, J(L_0^{-1})'\, X(s)' .$$

Proof. Let $H = \begin{bmatrix} Y \\ \dot{Y} \end{bmatrix}$; so that $H' = (Y', \dot{Y}')$, H, L and J are all $2n \times 2n$-matrices. Conditions (59) may be written in the abbreviated form

(62) $$L_0 H'_0 = J ,$$

where the subscript 0 means that the matrix in question is evaluated for $t = 0$. Since the system is self-adjoint, we see from (44) and (58) that $\ddot{Y}(t) = A(t)Y(t) + B(t)\dot{Y}(t)$. From (39) it follows that the columns of Y must be linear combinations of the columns of X. In other words there exists a $2n \times 2n$-matrix Q of constants such that

(63) $$Y(t) = X(t)Q$$

and hence $\dot{Y}(t) = \dot{X}(t)Q$. The last two written equations may be written more briefly as $H(T) = L(t)Q$. In particular, on setting $t = 0$, we have $H_0 = L_0 Q$.

Therefore $Q = L_0^{-1} H_0$. Inserting this into (63), we find that $Y(t) = X(t)L_0^{-1} H_0$, while from (62) we have $H'_0 = L_0^{-1} J$. Hence $H_0 = J'(L_0^{-1})'$. On inserting this expression for H_0 in the last formula for $Y(t)$, we see that (60) has been proved. And (61) now follows from (60) and (41).

In order to apply Theorem 5 to the system (30), we take $B(t) \equiv 0$, so that the condition (58) for self adjointness reduces to the requirement that $A(t)$ be symmetric. It actually is symmetric not only in the case of the variational equations of Keplerian motion but also in the case of the variational equations connected with any solution of a conservative holonomic dynamical system. In fact the matrix A in such a case is merely the negative of the Hessian matrix of the potential function with the given solution inserted, assuming that the coordinates are chosen in such a manner that the kinetic energy is given in the form $(1/2)(\dot{x} \cdot \dot{x})$.

V. THE Θ-MATRIX FOR ELLIPTIC MOTION

We now proceed to the consideration of the example alluded to at the end of Section III, restricting attention to the elliptic case. The matrix X of Section IV is now, in this special case, to be represented by the matrix W

of Section II. Hence the matrix $L(0)$ is the matrix $\begin{bmatrix} W(0) \\ \dot{W}(0) \end{bmatrix}$ and the inverse $L(0)^{-1}$ is accordingly the matrix of formula (25). Using this and various other formulas of Section II, the four elements of the matrix $\Theta(t, s)$ were calculated from the formula (61). We give only a brief outline of the tedious but straight forward calculation by mentioning only the following three intermediate formulas. First, by carrying out the indicated matrix multiplications, we have

$$L(0)^{-1} J(L(0)^{-1})' = \begin{bmatrix} 0 & +2 & 0 & 0 \\ -2 & 0 & 0 & \tan\lambda \\ 0 & 0 & 0 & -\sec\lambda \\ 0 & -\tan\lambda & \sec\lambda & 0 \end{bmatrix}$$

Secondly by using the identities of Section II, it is not difficult to establish the two following identities:

$$\psi \sec\lambda + \dot{\phi} \tan\lambda = -\psi\dot{\psi} \csc\lambda$$

$$\dot{\psi} \tan\lambda - \phi \sec\lambda = 1 + \phi\dot{\psi} \csc\lambda,$$

which are useful to eliminate the singularity at $\lambda = \pi/2$.

Let $\theta_{ij}(t, s)$ be the element in the $i^{\underline{th}}$ row and $j^{\underline{th}}$ column of $\Theta(t, s)$. Then we present our final results in the form:

$$\theta_{11}(t,s) = 3(s-t)\dot{\phi}(t)\dot{\phi}(s) + 2[\phi(t)\dot{\phi}(s) - \dot{\phi}(t)\phi(s)] - [\psi(t)\dot{\psi}(t)\phi_\lambda(s) - \psi(s)\dot{\psi}(s)\phi_\lambda(t)]\csc\lambda$$

$$\begin{aligned} \theta_{21}(t,s) = {} & 3(s-t)\dot{\psi}(t)\dot{\phi}(s) + 2[\psi(t)\dot{\phi}(s) - \dot{\psi}(t)\phi(s)] \\ & + [\phi(t)\dot{\psi}(t)\phi_\lambda(s) + \psi(s)\dot{\psi}(s)\psi_\lambda(t)] \csc\lambda + \phi_\lambda(s) \end{aligned}$$

$$\begin{aligned} \theta_{12}(t,s) = {} & 3(s-t)\dot{\phi}(t)\dot{\psi}(s) + 2[\phi(t)\dot{\psi}(s) - \dot{\phi}(t)\psi(s)] \\ & - [\psi(t)\dot{\psi}(t)\psi_\lambda(s) + \phi(s)\dot{\psi}(s)\phi_\lambda(t)] \csc\lambda - \phi_\lambda(t) \end{aligned}$$

$$\theta_{22}(t,s) = 3(s-t)\dot{\psi}(t)\dot{\psi}(s) + 2[\psi(t)\dot{\psi}(s) - \dot{\psi}(t)\psi(s)]$$
$$+ [\phi(t)\dot{\psi}(t)\psi_\lambda(s) - \phi(s)\dot{\psi}(s)\psi_\lambda(t)] \csc\lambda + [\psi_\lambda(s) - \psi_\lambda(t)] .$$

VI. ECCENTRICITY CHANGING TRANSFORMATIONS

In Section I we used the fact that the Keplerian differential equations were invariant under the scale, autonomous, and rotation transformation groups. These make it possible to pass from any trajectory to any other trajectory with the same eccentricity. The question naturally arises as to whether the equations are also invariant under a transformation group enabling a passage from any trajectory to another trajectory with different eccentricity. The answer appears to be in the affirmative provided that we allow a differential transformation on the time t instead of the simple transformation of the form $t' = P(t, x)$ as indicated in the introduction. This proviso causes difficulty in the use of such transformations, but nevertheless we present the theory in the hope that an application for it may be found at some future time.

Because of the rotation groups it is obviously permissible to confine attention to the planar problem, and indeed to orbits whose major axes lie along the x-axis of coordinates. We thus find it sufficient to consider the transformation $T(a, b)$, depending on the parameters a and b, defined by the equations,

$$(64)\qquad \begin{aligned} \xi &= x\sqrt{a^2 + b^2} + a\sqrt{x^2 + y^2} \\ \eta &= b\,y \qquad\qquad b \neq 0 . \end{aligned}$$

These transformations are easily seen to form a commutative group, since $T(0, 1)$ is the identity transformation, and since $T(a, b)^{-1} = T(-ab^{-2}, b^{-1})$, and $T(a, b)\,T(\alpha, \beta) = T(a\sqrt{\alpha^2 + \beta^2} + \alpha\sqrt{a^2 + b^2},\ b\beta)$.

These transformations all leave the origin invariant. They also leave the x-axis invariant. But, it is possible to pass from any point (x, y) with $y \neq 0$ to any other point (ξ, η) with $\eta \neq 0$. We have only to take

$$(65)\qquad a = \frac{\xi\sqrt{x^2 + y^2} - x\sqrt{\xi^2 + \eta^2}}{y^2}, \quad b = \frac{\eta}{y}$$

It is also possible to pass from any point $(x, 0)$ with $x \neq 0$ to any other point $(\xi, 0)$, with $\xi \neq 0$ and $\operatorname{sgn}\xi = \operatorname{sgn} x$.

Since it is proved easily from (64) that

(66) $$\sqrt{\xi^2 + \eta^2} = \sqrt{(a^2 + b^2)(x^2 + y^2)} + ax$$

it is seen (with $a \neq 0$) that the ellipse with eccentricity $|a|(a^2 + b^2)^{-\frac{1}{2}}$, focus at the origin, and directrix $x = a^{-1}$, is carried by the transformation into the unit circle $\xi^2 + \eta^2 = 1$. It can be seen further, restricting attention to conics with focus at the origin and directrices parallel to the y-axis, that a conic with eccentricity e is carried over by $T(a, b)^{-1}$ into a conic with eccentricity

$$\frac{e\sqrt{a^2 + b^2} + a}{\sqrt{a^2 + b^2} + ae} .$$

It follows that ellipses are carried into ellipses, parabolas into parabolas, and hyperbolas into hyperbolas. Except in the case of the parabola the eccentricity is always changed by any transformation of the group for which $a \neq 0$. Any ellipse can be carried into any other ellipse, any parabola into any other parabola, and any hyperbola into any other hyperbola. That the transformation carries any conic with focus at the origin (but with directrix not necessarily parallel to the y-axis) into another conic with focus at the origin is more difficult to see; but it is one of the conclusions that can be drawn from the following discussion of the Keplerian differential equations.

We now examine the effect of the transformation (64), where a and b are regarded as constants while $x = x(t)$ and $y = y(t)$ are functions satisfying the Keplerian equations $\ddot{x} = -x r^{-3}$, $\ddot{y} = -y r^{-3}$, where $r^2 = x^2 + y^2$. Thus ξ and η are also to be regarded as functions of t. We also introduce $\rho^2 = \xi^2 + \eta^2$ and another variable σ defined uniquely by the requirements that $d\sigma/dt = \rho/r$ and that $\sigma = 0$ when $t = 0$. Thus σ is a function of t, but we can also solve for t in terms of σ and use the latter for the independent variable. In this way ξ and η can be regarded as functions of σ and we write $\xi' = d\xi/d\sigma = (d\xi/dt)(r/\rho)$, with similar formulas for η. We wish to find differential equations (similar to the Keplerian differential equations) to be satisfied by $\xi(\sigma)$ and $\eta(\sigma)$.

It is known that $x\dot{y} - \dot{x}y = c$ is independent of t (and hence also of

σ). This is also true of $c\dot{y} - x\,r^{-1} = A$ and $c\dot{x} + y\,r^{-1} = B$.

A straight forward calculation shows that the quantity $\gamma = \xi\eta' - \xi'\eta$ is also independent of t (and hence of σ). In fact it shows that $\gamma = bc$. Still further, if we introduce the quantities, $Q = (a^2 + b^2)^{1/2} - aA$, $M = bB$, and $N = A(a^2 + b^2)^{1/2}$, all of which are independent of t and σ, our straight forward calculation shows that $\gamma\xi' + Q\eta\rho^{-1} = M$ and $\gamma\eta' - Q\xi\rho^{-1} = N$. Multiplying the first of these two equations by η and the second by ξ, we find after a subtraction that $\gamma(\xi\eta' - \xi'\eta) = Q\rho + N\xi - M\eta$. Hence $\gamma^2 - Q\rho = N\xi - M\eta$. It is also easy to prove that

$$\gamma^2(\xi'^2 + \eta'^2) = (M^2 + N^2) + 2Q\rho^{-1}(N\xi - M\eta) + Q^2$$

$$= (M^2 + N^2) + 2Q\rho^{-1}(\gamma^2 - Q\rho) + Q^2$$

$$= M^2 + N^2 - Q^2 + \frac{2Q\gamma^2}{\rho} .$$

Thus we have established the following two relations

$$\xi\eta' - \xi'\eta = bc$$

$$\frac{1}{2}(\xi'^2 + \eta'^2) = Q\rho^{-1} + (M^2 + N^2 - Q^2)/(2\gamma^2) .$$

If we now differentiate these two equations with respect to σ and solve for the second derivatives ξ'' and η'', we find that $\xi'' = -Q\xi\rho^{-3}$, $\eta'' = -Q\eta\rho^{-3}$. Since Q is a constant, these equations are already in Keplerian form; so that it is clear that the point $(\xi(\sigma), \eta(\sigma))$ describes an orbit which must be a conic section, and, if it is a non-degenerate conic, one focus must be at the origin.

Nevertheless Q need not be unity. Hence in order to achieve the invariance of the original equations $\ddot{x} = -x\,r^{-3}$, $\ddot{y} = -y\,r^{-3}$, we make another modification of the independent variable. Let τ be defined in such wise that $d\tau/d\sigma = Q^{1/2}$, with $\tau = 0$ when $\sigma = 0$. Then using τ as the independent variable, we find that $d^2\xi/d\tau^2 = -\xi\rho^{-3}$ and $d^2\eta/d\tau^2 = -\eta\rho^{-3}$, where the relation between τ and t is evidently

$$\frac{d\tau}{dt} = \frac{\rho}{r} Q^{1/2} = \frac{\rho}{r} [(a^2 + b^2)^{1/2} - aA]^{1/2}$$

Although A is a constant of the motion, it may change from one motion to another. Hence it has to be considered as an abbreviation for

$(x\dot{y} - \dot{x}y)\dot{y} - x(x^2 + y^2)^{-1/2}$.

We conclude therefore that there exists a function $U(x, y, \dot{x}, \dot{y}, a, b)$ such that the transformation

$$\xi = x\sqrt{a^2 + b^2} + a\sqrt{x^2 + y^2}$$

$$\eta = b\,y \qquad\qquad b \neq 0$$

$$d\tau/dt = U(x, y, \dot{x}, \dot{y}, a, b)$$

transforms the Keplerian differential system $\ddot{x} = -x(x^2 + y^2)^{-3/2}$, $\ddot{y} = -y(x^2 + y^2)^{-3/2}$ into the Keplerian system,

$$\frac{d^2\xi}{d\tau^2} = -\frac{\xi}{\rho^3}, \quad \frac{d^2\eta}{d\tau^2} = -\frac{\eta}{\rho^3}.$$

There is a question about the sign of the quantity Q; for, if it is negative, our change in the time variable through $d\tau/d\sigma = Q^{1/2}$ would be imaginary. We have not investigated this situation, except to note that Q is necessarily positive when the motion $(x(t), y(t))$ is either elliptic or parabolic. For it is well known that the constants A and B, introduced above, are related to the eccentricity e through the equation

$$e^2 = A^2 + B^2 .$$

Hence $|A| \leq e \leq 1$ in either the elliptic or parabolic case, and since we must always take $b \neq 0$, we see immediately that

$$Q = (a^2 + b^2)^{1/2} - aA > 0 .$$

COMMENTS ON THE SUNDMAN INEQUALITY

By D. C. Lewis
Control Research Associates
Baltimore, Maryland

NASA Contract 12-93

COMMENTS ON THE SUNDMAN INEQUALITY

By Daniel C. Lewis

Professor of Applied Mathematics

The Johns Hopkins University

Baltimore, Maryland

SUMMARY

A new proof of the Sundman inequality with refinements, and its utility in the introduction of a monotonically increasing angular variable for the Lagrangian inertial radius in the n-body problem. A discussion of the possible amount of information derivable from these considerations by studying them in the light of the integrable two-body problem.

INTRODUCTION

In this paper we give a proof of the well known Sundman inequality which is far superior in generality, precision, and elegance to any which we have hitherto seen. The additional terms to be introduced in order to turn the inequality into an equality are explicitly exhibited in a reasonably simple form. Thus, by including some or all of these terms, we achieve the superior precision noted above. All these results are based on some simple theorems in the field of classical vector analysis. In fact the essence of the Sundman inequality appears in a more general setting than that of the n-body problem.

The last sentence needs to be emphasized despite the fact that the only applications of the inequality have been to the theory of the n-body problem, and despite also the fact that in this very paper we have ventured (through the Sundman inequality) a further modest contribution to this theory. Namely we have shown how to introduce a monotonically increasing "angular" variable $\theta(t)$ with certain interesting properties. Upon setting $X = R\cos\theta$ and $Y = R\sin\theta$, where R is the Lagrangian inertial radius, we investigate differential equations to be satisfied by X and Y. These equations are shown to contain enough information essentially to solve the problem completely when $n = 2$. Naturally the situation for $n > 2$ is very different, since the more complicated cases need much more information for their complete solution. But it may be reasonably hoped that the same information which solves the two-body problem completely may at least yield some interesting qualitative results for the n-body problem.

I. PRELIMINARY THEOREMS IN VECTOR ANALYSIS

In the sequel the inner or scalar product of two 3-vectors a and b is denoted either by (ab), without a dot, or by $a \cdot b$, without the parentheses. The outer or vector product will be denoted by $(a \times b)$.

Theorem 1. Let a, b, c, d be any four vectors in three dimensional space. Then the scalar $W(a, b, c, d) = (aa)(dd) - 2(ac)(bd) - 2(ab)(cd) + 2(ad)(bc) + (bb)(cc)$ is not negative.

Proof. The theorem is certainly true if $d = 0$; for then $W = (bb)(cc) \geq 0$. Hence we may fix attention on the case $d \neq 0$. Then, regarding b, c, and d as constant vectors, we seek to minimize W by varying a. Evidently the gradient of W with respect to the components of a is the vector

$$\partial W/\partial a = 2a(dd) - 2c(bd) - 2b(cd) + 2d(bc) ,$$

while the hessian matrix of W with respect to the components of a is the positive definite matrix, $2(dd)I$, where I is the 3×3 identity matrix. It follows that W assumes its minimum value when a is such that it makes the gradiant of W vanish, namely, when

$$a = (dd)^{-1} [b(cd) + c(bd) - d(bc)] .$$

If we substitute this expression for a in the formula for W, we find by a routine calculation that the minimum value of W turns out to be

$$(dd)^{-1} \begin{vmatrix} (bb) & (bc) & (bd) \\ (cb) & (cc) & (cd) \\ (db) & (dc) & (dd) \end{vmatrix}$$

We have here the determinant of a certain type of symmetric matrix which is well known to be positive if the vectors b, c, and d are linearly independent and, otherwise, it is positive semi-definite. Cf. Hardy, Littlewood, and Polya, Inequalities, Cambridge University Press (1934), p. 16. In either case, the minimum value of the determinant (and hence of W) must be non-negative and thus the theorem is proved.

Theorem 2. Let $u_1, u_2, \ldots, u_n$ be n vectors in ordinary three dimensional space and let $v_1, v_2, \ldots, v_n$ be another set of n vectors. Then the following formula is identically satisfied

$$\left(\sum_{i=1}^{n}(u_iu_i)\right)\left(\sum_{i=1}^{n}(v_iv_i)\right)-\left(\sum_{i=1}^{n}(u_iv_i)\right)^2-\left(\sum_{i=1}^{n}(u_i\times v_i)\right)\cdot\left(\sum_{i=1}^{n}(u_i\times v_i)\right)=\frac{1}{2}\sum_{i,j=1}^{n}W(u_i,u_j,v_i,v_j)$$

where the W function is defined by the formula of Theorem 1.

Proof. Evidently the left member of the last formula may be written in the form

$$\sum_{i,j=1}^{n}\left[(u_iu_i)(v_jv_j)-(u_iv_i)(u_jv_j)-(u_i\times v_i)\cdot(u_j\times v_j)\right]$$

By a known formula of vector analysis, we know that

$$(u_i\times v_i)\cdot(u_j\times v_j)=(u_iu_j)(v_iv_j)-(u_iv_j)(v_iu_j),$$

while a simple manipulation of the summation indices i and j shows that

$$\sum_{i,j=1}^{n}(u_iu_i)(v_jv_j)=\sum_{i,j=1}^{n}\left[\frac{1}{2}(u_iu_i)(v_jv_j)+\frac{1}{2}(u_ju_j)(v_iv_i)\right]$$

Hence the left member of our formula can be transformed into

$$\sum_{i,j=1}^{n}\left[\frac{1}{2}(u_iu_i)(v_jv_j)-(u_iv_i)(u_jv_j)-(u_iu_j)(v_iv_j)+(u_iv_j)(u_jv_i)+\frac{1}{2}(u_ju_j)(v_iv_i)\right]$$

which, by the formula for W, given in Theorem 1, can be written

$$\sum_{i,j=1}^{n}\frac{1}{2}W(u_i,u_j,v_i,v_j)$$

and this finishes the proof.

As a corollary to Theorem 1 and 2, we obtain the inequality

$$(1)\qquad \left(\sum_{i=1}^{n}(u_iu_i)\right)\left(\sum_{i=1}^{n}(v_iv_i)\right)\geqq\left(\sum_{i=1}^{n}(u_iv_i)\right)^2+\left(\sum_{i=1}^{n}(u_i\times v_i)\right)\cdot\left(\sum_{i=1}^{n}(u_i\times v_i)\right)$$

no matter what the vectors $u_1,\ldots,u_n$, $v_1,\ldots,v_n$ may be.

Theorem 3. The scalar $W(a, b, c, d)$ of Theorem 1 vanishes whenever the vectors c and d are proportional to the vectors a and b, in the sense that there exists a scalar k such that $b = ka$ and $d = kc$.

The proof of this theorem is left to the reader since it is an almost obvious consequence of the formula for $W(a, b, c, d)$.

II. THE SUNDMAN INEQUALITY

In the application to the n-body problem with the n masses $m_1, m_2, \ldots, m_n$, we let the position of the mass m_i at time t be determined relative to a given coördinate frame by the position vector r_i with components x_i, y_i, z_i. The velocity of m_i is then $\dot{r}_i$ with components $\dot{x}_i, \dot{y}_i, \dot{z}_i$, the dot denoting differentiation with respect to t. The Lagrange inertial radius R is defined by

$$(2) \qquad R^2 = \sum_{i=1}^{n} m_i(r_i r_i) .$$

The kinetic energy T is given by

$$(3) \qquad T = \frac{1}{2} \sum_{i=1}^{n} m_i(\dot{r}_i \dot{r}_i) .$$

The angular momentum vector ϕ is evidently

$$(4) \qquad \phi = \sum_{i=1}^{n} m_i(r_i \times \dot{r}_i) .$$

Also, by differentiating (2) we have

$$(5) \qquad R\dot{R} = \sum_{i=1}^{n} m_i(r_i \dot{r}_i) .$$

If now we let $u_i = m_i^{1/2} r_i$ and $v_i = m_i^{1/2} \dot{r}_i$, we see that (2), (3), (4) and (5) become

$$R^2 = \sum_{i=1}^{n} (u_i u_i)$$

$$2T = \sum_{i=1}^{n} (v_i v_i)$$

$$\phi = \sum_{i=1}^{n} (u_i \times v_i)$$

$$R\dot{R} = \sum_{i=1}^{n} (u_i v_i)$$

Thus, substituting in our inequality (1), we obtain one form of the famous Sundman inequality, namely,

$$2R^2T \geq R^2 \dot{R}^2 + (\phi \cdot \phi) \; . \tag{6}$$

We may also write this in the form

$$\dot{R}^2 + \frac{f^2 + Q}{R^2} = 2T \tag{7}$$

where $f = (\phi \cdot \phi)^{1/2}$ is the magnitude of the angular momentum and Q is a positive scalar. Actually, by using the identity of Theorem 2 instead of the inequality (1), we easily derive the following explicit formula for Q

$$\begin{aligned} Q &= \frac{1}{2} \sum_{i,j=1}^{n} W(u_i, u_j, v_i, v_j) \\ &= \frac{1}{2} \sum_{i,j=1}^{n} W(m_i^{1/2} r_i, m_j^{1/2} r_j, m_i^{1/2} \dot{r}_i, m_j^{1/2} \dot{r}_j) \; . \end{aligned} \tag{8}$$

So far we have made no use of the equations of motion for the n-body problem. The Sundman formulas (6) and (7), with Q defined by (8), hold for any system of n masses moving around in space with arbitrary velocities. As a consequence of the equations of motion, we know that we have ten first integrals, from which it follows that

$$f = \text{constant} \tag{9}$$

and that there exists a constant K (the negative of the energy constant) such that

$$\frac{d^2R^2}{dt^2} + 2K = 2T \; . \tag{10}$$

This last relation is one form of the well known Lagrange identity. On using (10) to eliminate the T in (7), we obtain

(11) $$\dot{R}^2 + 2R\ddot{R} + 2K \geq f^2 R^{-2} ,$$

which is the form of Sundman's inequality used in our previous report entitled Rejection to Infinity in the Problem of Three Bodies when the Total Energy is Negative.

III. AN ANGULAR VARIABLE FOR R

It is convenient to write (7) in the form

(12) $$\dot{R}^2 + \frac{p^2}{R^2} = 2T$$

where $p^2 = f^2 + Q \geq f^2$. We then define a monotonically increasing function $\theta(t)$ by means of

(13) $$\theta(t) = \theta_0 + \int_0^t p(s)\ R(s)^{-2} ds$$

where θ_0 is an arbitrary constant. Hence $R(t)\ \dot{\theta}(t) = p(t)\ R(t)^{-1}$ so that (12) becomes

(14) $$\dot{R}^2 + R^2\ \dot{\theta}^2 = 2T$$

while, of course, we also have

(15) $$R^2\ \dot{\theta} = p \geq f \geq 0$$

Now let

(16) $$X(t) = R \cos \theta \qquad Y(t) = R \sin \theta$$

Then we easily find that

(17) $$\begin{aligned} X^2 + Y^2 &= R^2 \\ \dot{X}^2 + \dot{Y}^2 &= \dot{R}^2 + R^2\ \dot{\theta}^2 = 2T \\ X\ \dot{Y} - \dot{X}\ Y &= R^2\ \dot{\theta} \end{aligned}$$

Hence

$$\frac{1}{2}(\dot{X}^2 + \dot{Y}^2) = T \quad \text{and} \quad X\dot{Y} - \dot{X}Y = p \ . \tag{18}$$

From (10) and (12), we have

$$\dot{R}^2 + p^2 R^{-2} = \frac{d^2R^2}{dt^2} - 2h,$$

where $h = -K$ is the energy constant. From this we find that

$$\frac{d}{dt}(R\dot{R}^2 - 2h\,R + \frac{p^2}{R}) = \frac{2p\,\dot{p}}{R} \ .$$

Integrating we get

$$R\dot{R}^2 - 2h\,R + p^2\,R^{-1} = c(t)$$

where

$$c(t) = R_0\dot{R}_0^{\,2} - 2h\,R_0 + p_0^{\,2}\,R_0^{\,-1} + \int_0^t 2R(s)^{-1}\,p(s)\,\dot{p}(s)ds \ .$$

From (15) we now obtain

$$\dot{R}^2 + R^2\,\dot{\theta}^2 = c\,R^{-1} + 2h \ .$$

So from (17) we have

$$\dot{X}^2 + \dot{Y}^2 = c\,R^{-1} + 2h$$

$$X\dot{Y} - \dot{X}Y = p$$

Differentiating these last two equations we find that

$$\dot{X}\ddot{X} + \dot{Y}\ddot{Y} = -\frac{c}{2}\frac{\dot{R}}{R^2} + \frac{p\,\dot{p}}{R^2} \tag{19a}$$

$$-Y\ddot{X} + X\ddot{Y} = \dot{p} \ . \tag{19b}$$

In deriving (19a) we, of course, use the obvious fact that $\dot{c} = 2R^{-1}\,p\,\dot{p}$. We now solve (19) for $\ddot{X}$ and $\ddot{Y}$, using the fact, derived from (17) that

$R\dot{R} = X\dot{X} + Y\dot{Y}$. In this manner we obtain the pair of equations

$$\text{(20)}\qquad \begin{aligned} \ddot{X} &= \frac{-\mu X}{R^3} + \frac{X p \dot{p}}{R^3 \dot{R}} - \frac{\dot{p}\dot{Y}}{R\dot{R}} \\ \ddot{Y} &= \frac{-\mu Y}{R^3} + \frac{Y p \dot{p}}{R^3 \dot{R}} + \frac{\dot{p}\dot{X}}{R\dot{R}} \end{aligned}$$

where

$$\mu = \frac{1}{2} c = \frac{1}{2}\left[R_0 \dot{R}_0^{\,2} - 2h R_0 + \frac{p_0^{\,2}}{R_0}\right] + \int_0^t \frac{p(s)\,\dot{p}(s)}{R(s)}\, ds \; .$$

Thus, if $\dot{p}$ is small relative to $\dot{R}$ and R, then X and Y very nearly satisfy the equations for Keplerian motion.

We now show that the information contained in (20), in the case of the two body problem, is equivalent to the reduction of the latter to Keplerian motion.

Taking the origin at the center of gravity of the system, we have

$$\text{(21)}\qquad m_1 r_1 + m_2 r_2 = 0, \qquad m_1 \dot{r}_1 + m_2 \dot{r}_2 = 0 \; .$$

Hence the vectors $m_1^{1/2}\,\dot{r}_1$ and $m_2^{1/2}\,\dot{r}_2$ are proportional to the vectors $m_1^{1/2}\,r_1$ and $m_2^{1/2}\,r_2$. Hence from Theorem 3 and equation (8) we see that $Q = 0$; and therefore $p = f$, which is a constant. Hence equations (2) reduce to

$$\text{(22)}\qquad \ddot{X} = -\frac{\mu X}{R^3} \; , \qquad \ddot{Y} = \frac{-\mu Y}{R^3}$$

where μ is the constant

$$\mu = \frac{1}{2}\left[R_0 \dot{R}_0^{\,2} - 2h R_0 + f^2 R_0^{-1}\right] .$$

So far, we know only that μ is a constant along each motion; but it might differ for different motions. We show now, however, that the latter is not the case. Indeed, we can show that μ is a simple function of m_1 and m_2 only. To this end we write

$$\mu = \frac{R}{2}\left[\dot{R}^2 + f^2 R^{-2} - 2h\right] ,$$

dropping the subscript 0, which is not important, since the right hand member of the last written equation is almost obviously a constant along each motion.

Now by (15) we have $f R^{-1} = p R^{-1} = R \dot{\theta}$, so that

$$\mu = \frac{R}{2} [\dot{R}^2 + R^2 \dot{\theta}^2 - 2h] .$$

It now follows from (17) that $\mu = R(T - h)$. Since h is the total energy of the system, $h - T$ must be the potential energy of the system, which is, of course, $- m_1 m_2 \rho^{-1}$, where ρ is the distance between m_1 and m_2. It follows that

$$\mu = m_1 m_2 R \rho^{-1} \tag{23}$$

By definition of R and by (21) we have

$$R^2 = m_1 |r_1|^2 + m_2 |r_2|^2 = |r_1|^2 (m_1/m_2)(m_1 + m_2) . \tag{24}$$

Similarly

$$\rho = |r_1| + |r_2| = |r_1| m_2^{-1} (m_1 + m_2) . \tag{25}$$

Hence, from (23), we have

$$\mu = \left(\frac{m_1^3 m_2^3}{m_1 + m_2}\right)^{1/2}$$

If we orient our coördinate frame of reference so that the positive z-axis has the direction of the angular momentum vector, the motion takes place in the xy-plane and we obtain the equations

$$\ddot{\xi} = - \frac{(m_1 + m_2)\xi}{\rho^3} , \ddot{\eta} = - \frac{(m_1 + m_2)\eta}{\rho^3} \tag{26}$$

for the relative motion of the two bodies, where ξ and η are the coordinates of m_1 (say) relative to m_2. The ratio of $\rho = (\xi^2 + \eta^2)^{1/2}$ to R is seen from (24) and (25) to be $(m_1 + m_2)^{1/2}/(m_1 m_2)^{1/2}$. By a discussion of the angular momentum it may be shown that $\tan^{-1}(\eta/\xi) - \theta$ is a constant; and it is also not hard to prove that, if the axes are oriented so that this constant is zero, then the ratio ξ to X (and of η to Y) must be a constant and equal to the ratio of ρ to R. It is then easy to see that (22) and (26) are equivalent.

APPROXIMATE DECOUPLING IN THE n-BODY PROBLEM

By D. C. Lewis
Control Research Associates
Baltimore, Maryland

NASA Contract NAS 12-93

APPROXIMATE DECOUPLING IN THE n-BODY PROBLEM

by Daniel C. Lewis, Jr.

Professor of Applied Mathematics

The Johns Hopkins University

Baltimore, Maryland

SUMMARY

The intuitive aspects of approximate decoupling of an n-body problem as exemplified in the solar system. Formalization of approximate decoupling in terms of the limiting values of certain parameters contained in the equations. Partial decoupling versus complete decoupling. Details in the approximate decoupling of an n-body problem into a k-body problem $(k < n)$ and an $(n-k+1)$-body problem in the situation where k of the bodies are relatively close to each other compared to the other mutual distances and where also the k bodies have small masses compared to the other $(n-k)$ bodies. The Hill equations for the motion of the moon as a special case of partial decoupling in the above situation with $n = 3$ and $k = 2$.

Theory of so-called quasi-first integrals which appear in the partially decoupled system and which are derived from first integrals of the unreduced system. Examples in the reduced and restricted problems of three bodies.

A method for appraising the validity of an approximate decoupling (either partial or complete) based on the use of Lipschitz constants. The hyperbolic cosine appraisal for comparing the solutions of two systems of second order equations.

There is a problem in applying this appraisal to the n-body equations because of the collision singularities. If these singularities are excluded by restricting attention to a region in configuration space in which each mutual distance between pairs of bodies is not less than a fixed positive number assigned to the pair in question, the resulting region is not convex; and so there is still a problem in the estimation of the Lipschitz constant in terms of bounds for partial derivatives. This problem is solved for the appropriate functions occurring in the n-body problem by proving that in this special case the non-convexity of the region may be ignored.

INTRODUCTION

One of the remarkable features of an n-body problem, of the type afforded by the solar system, is the ease with which the system can be approximately decoupled. Thus the motions of the planets about the sun are usually treated as perturbed Keplerian motions, as are also the motions of satellites about a planet. This approximate decoupling is possible when some of the masses are very small compared to others and/or when some of the mutual distances between the bodies are very large compared to the other distances.

If we introduce parameters into the system in a suitable way, the ensemble of the decoupled systems may be regarded as the limiting case of the original n-body system when one or more of these parameters approach certain limiting values. For instance, a three body problem in which one of the masses is very large compared with the other two is partially decoupled into one Keplerian motion and one restricted (or reduced) problem of three bodies when one of the smaller masses approaches zero; and it is completely decoupled into two independent Keplerian motions when both of the two smaller masses approach zero. This is the type of approximate decoupling observed in planetary theory. Another type of decoupling occurs in linear theory where the distance between the two smaller bodies is small relative to their distances from the largest mass. Again we may take one of the parameters to be the smallest mass and we get the same partial decoupling as before, when this parameter tends to zero. But, for the other parameter, we take something which tends to zero with the ratios of the smallest distance to the two larger distances, and simultaneously modifies the time scale. There are several ways of doing this; but, if it is done properly, the three body problem will again be approximately decoupled into two Keplerian motions, one for the motion of the intermediate mass (the earth, say) about the largest mass (the sun) and the other for the motion of the smallest mass (the moon) about the intermediate mass (the earth). But the time intervals for the validity of these approximations may be vastly different. Considering only the situation where the two above Keplerian motions are elliptic, the approximations may be reasonably valid in each case for approximately the same number of periods, but the period associated with the smallest mass may be but a small fraction of the period associated with the intermediate mass.

In this paper we wish first to consider a method for the approximate partial decoupling of the n-body problem into an $(n-k+1)$-body problem and a modified k-body problem $(k<n)$ when the k bodies are relatively close to each other compared to the other mutual distances. We also assume that these k bodies are small compared to the other $n-k$ bodies. When $n=3$, $k=2$, this modified k-body problem is the problem formulated by the Hill equations, at least if the $(n-k+1)$-body problem (i. e. in this case, the 2-body problem) is given a circular solution. By introducing a second parameter and proceeding again to a limiting case we may achieve a complete approximate decoupling.

Secondly we shall initiate a general theory of so-called quasi-first integrals which arise, in the partially decoupled system, from the first integrals of the original system when the parameter takes on its limiting value, say 0.

Thirdly, we give a general but rather crude method for estimating the error consequent upon using the decoupled or partially decoupled systems in place of the original system.

I. DECOUPLING OF THE n-BODY PROBLEM

I.1. Partial Decoupling

The original equations of motion are written as follows in terms of position vectors q for the n-bodies (with masses $m_1, m_2, \ldots, m_n$):

$$(1) \qquad m_i \ddot{q}_i = U_{q_i} \qquad i = 1, \ldots, k$$

$$(2) \qquad m_\alpha \ddot{q}_\alpha = U_{q_\alpha} \qquad \alpha = k+1, \ldots, n$$

where the dots represent differentiations with respect to the time t, where the subscript q's in (1) and (2) indicate the gradient of U with respect to the relevant q, and where U itself is defined by the following formulas

$$(3) \qquad \begin{aligned} U &= U^* + V + W \\ U^* &= \sum_{i<j} \frac{m_i m_j}{|q_j - q_i|} && j = 1, \ldots, k \\ V &= \sum_{\alpha=k+1}^{n} \sum_{i=1}^{k} \frac{m_\alpha m_i}{|q_\alpha - q_i|} \\ W &= \sum_{\alpha<\beta} \frac{m_\alpha m_\beta}{|q_\alpha - q_\beta|} && \beta = k+1, \ldots, n. \end{aligned}$$

Setting $\mu = m_1 + m_2 + \ldots + m_k$, the center of gravity of these k bodies is given by

$$(4) \qquad q_0 = \mu^{-1} \sum_{j=1}^{k} m_j q_j$$

Letting $r_i = q_i - q_0$, $r_\alpha = q_\alpha - q_0$, we see that $r_j - r_i = q_j - q_i$,

$r_\alpha - r_i = q_\alpha - q_i$, $r_\beta - r_\alpha = q_\beta - q_\alpha$. Hence the formulas for U^*, V, W remain valid when the q's are replaced by the r's. It is also clear that $\sum_{i=1}^{k} m_i r_i = \Sigma m_i(q_i - q_0) = \Sigma m_i q_i - \Sigma m_i q_0 = \Sigma m_i q_i - \mu q_0 = 0$ by (4). Hence

$$\sum_{i=1}^{k} m_i r_i = 0 . \tag{5}$$

Evidently $m_i\ddot{r}_i = m_i\ddot{q}_i - m_i\ddot{q}_0 = m_i\ddot{q}_i - m_i\mu^{-1}\sum_{j=1}^{k} m_j\ddot{q}_j$

$$= U_{r_i} - m_i\mu^{-1}\sum_{j=1}^{k} U_{r_j}$$

It is easy to see by direct calculation that

$$\sum_{\ell=1}^{k} U^*_{r_\ell} = 0 \quad \text{and} \quad \sum_{\beta=k+1}^{n} W_{r_\beta} = 0 . \tag{6}$$

Hence

$$m_i\ddot{r}_i = U^*_{r_i} + V_{r_i} - m_i\mu^{-1}\sum_{j=1}^{k} V_{r_j} . \tag{7}$$

Similarly, we find that

$$m_\alpha\ddot{r}_\alpha = V_{r_\alpha} + W_{r_\alpha} - m_\alpha\mu^{-1}\sum_{j=1}^{k} V_{r_j} \tag{8}$$

The equations of the n-body problem are easily seen to be invariant under a transformation which multiplies each distance by a parameter s provided that each mass is correspondingly multiplied by s^3. Hence, if $r_{k+1}, \ldots, r_n$ are large compared to $r_1, \ldots, r_k$, and if one wishes to exaggerate the comparative largeness of the former to the latter, it would appear desirable to multiply each mass by s^3 and each of the vectors $r_{k+1}, \ldots, r_n$ by s, but leave $r_1, \ldots, r_k$ untouched. One then examines the effect of allowing s to approach infinity. Carrying out these modifications on (8), and making use of the formulas of (3) for V and W, we obtain

$$m_\alpha s^4 \ddot{r}_\alpha = \sum_{i=1}^{k} \frac{s^6 m_i m_\alpha (r_i - s r_\alpha)}{|r_i - s r_\alpha|^3} + \sum_{\beta \neq \alpha} \frac{s^6 m_\alpha m_\beta (s r_\beta - s r_\alpha)}{|s r_\beta - s r_\alpha|^3}$$

$$- m_\alpha \mu^{-1} \sum_{j=1}^{k} \sum_{\beta=k+1}^{n} \frac{s^6 m_n m_\beta (s r_\beta - r_j)}{|r_j - s r_\beta|^3} .$$

Dividing by $s^4 m_\alpha$ and then letting $s \to \infty$, we obtain

$$\ddot{r}_\alpha = - \frac{\mu r_\alpha}{|r_\alpha|^3} + \sum_{\beta \neq \alpha} \frac{m_\beta (r_\beta - r_\alpha)}{|r_\beta - r_\alpha|^3} - \sum_{\beta=k+1}^{n} \frac{m_\beta r_\beta}{|r_\beta|^3} . \tag{9}$$

These are precisely the heliocentric equations of notion for $n-k+1$ bodies, namely the $n-k$ bodies $m_{k+1}, \ldots, m_n$ referred to a hypothetical sun with mass $\mu = m_1 + \ldots + m_k$, placed at the center of gravity of the bodies $m_1, m_2, \ldots, m_k$. Equations (9) thus approximately describe the notion of the $n-k$ bodies $m_{k+1}, \ldots, m_n$ and the center of mass of the k bodies $m_1, \ldots, m_k$, while neglecting $|r_1|, \ldots, |r_k|$ in comparison with $|r_{k+1}|, \ldots, |r_n|$, but not neglecting the total attraction of $m_1, \ldots, m_k$ on the other $n-k$ bodies.

If we also wish to neglect the masses $m_1, \ldots, m_k$ in comparison with $m_{k+1}, \ldots, m_n$, we follow the procedure as above except that we do not introduce the factor s^3 in connection with each of the k first masses, but only with the remaining masses. We then find, on letting $s \to \infty$, that

$$\ddot{r}_\alpha = \sum_{\beta \neq \alpha} \frac{m_\beta (r_\beta - r_\alpha)}{|r_\beta - r_\alpha|^3} - \sum_{\beta=k+1}^{n} \frac{m_\beta r_\beta}{|r_\beta|^3} , \tag{10}$$

which, of course, is the same as (9), except that the first term is omitted on the right hand side.

We now investigate the behavior of equations (7) under these operations, neglecting $m_1, \ldots, m_k$ in comparison with $m_{k+1}, \ldots, m_n$ as well as $r_1, \ldots, r_k$ in comparison with $r_{k+1}, \ldots, r_n$. The equations may be written after the $i^{\underline{th}}$ equation is divided by m_i in the following manner:

$$\ddot{r}_i = \sum_{j\neq i} \frac{m_j(r_j - r_i)}{|r_j - r_i|^3} + \sum_{\alpha=k+1}^{n} \frac{m_\alpha(r_\alpha - r_i)}{|r_\alpha - r_i|^3} - \mu^{-1} \sum_{j=1}^{k} \sum_{\alpha=k+1}^{n} \frac{m_j m_\alpha(r_\alpha - r_j)}{|r_\alpha - r_j|^3}$$

Hence multiplying r_α by s $(\alpha = k+1, \ldots, n)$ and m_α by s^3 and then letting $s \to \infty$, we get

$$\ddot{r}_i = \sum_{j\neq i} \frac{m_j(r_j - r_i)}{|r_j - r_i|^3} + \tag{11}$$

$$\lim_{s\to\infty} \left[\sum_{\alpha=k+1}^{n} \frac{s^3 m_\alpha(sr_\alpha - r_i)}{|sr_\alpha - r_i|^3} - \mu^{-1} \sum_{j=1}^{k} \sum_{\alpha=k+1}^{n} \frac{m_j s^3 m_\alpha(sr_\alpha - r_j)}{|sr_\alpha - r_j|^3}\right]$$

In order to evaluate the limit $s \to \infty$, we introduce $\sigma = 1/s$. We then find that

$$\frac{s^3 m_\alpha(sr_\alpha - r_i)}{|sr_\alpha - r_i|^3} - \mu^{-1} \sum_{j=1}^{k} \frac{m_j s^3 m_\alpha(sr_\alpha - r_j)}{|sr_\alpha - r_j|^3}$$

$$= \sigma^{-1} \left[\frac{m_\alpha(r_\alpha - \sigma r_i)}{|r_\alpha - \sigma r_i|^3} - \mu^{-1} \sum_{j=1}^{k} \frac{m_j m_\alpha(r_\alpha - \sigma r_j)}{|r_\alpha - \sigma r_j|^3}\right] .$$

The quantity in the bracket is now expanded in a power series in σ. It is obvious, in virtue of $\mu = \sum_{j=1}^{k} m_j$, that the constant term is zero, while the coefficient of σ is readily computed to be

$$-\frac{m_\alpha r_i}{r^3} + \frac{3m_\alpha r_\alpha(r_\alpha \cdot r_i)}{|r_\alpha|^5} .$$

In arriving at this result we use (5). It follows that (11) may be written in the form

$$\ddot{r}_i = \sum_{j\neq i} \frac{m_j(r_j - r_i)}{|r_j - r_i|^3} + \sum_{\alpha=k+1}^{n} \left(-\frac{m_\alpha r_i}{|r_\alpha|^3} + \frac{3m_\alpha r_\alpha(r_\alpha \cdot r_i)}{|r_\alpha|^5}\right) . \tag{12}$$

I.2. The Hill Example

We now indicate the manner in which the Hill equations are special cases of equations (12). We thus assume $n = 3$, $k = 2$, so that $k + 1 = n = 3$. We also restrict attention to the planar problem. Equation (10) reduces to

$$\ddot{r}_3 = \frac{-m_3 r_3}{|r_3|^3} \tag{13}$$

since the only value that α and β are allowed to take on is 3. This is, of course, just the system for Keplerian motion. Suppose we take $m_3 = 1$ and consider the particular solution $x_3 = \cos t$, $y_3 = \sin t$. So $|r_3| = 1$.

Hence equations (12) become

$$(14)\qquad \begin{aligned} \ddot{x}_1 &= \frac{m_2(x_2-x_1)}{[(x_2-x_1)^2 + (y_2-y_1)^2]^{3/2}} - x_1 + 3\cos t\,(x_1 \cos t + y_1 \sin t) \\ \ddot{y}_1 &= \frac{m_2(y_2-y_1)}{[(x_2-x_1)^2 + (y_2-y_1)^2]^{3/2}} - y_1 + 3\sin t\,(x_1 \cos t + y_1 \sin t) \end{aligned}$$

$$(15)\qquad \begin{aligned} \ddot{x}_2 &= \frac{m_1(x_1-x_2)}{[(x_1-x_2)^2 + (y_1-y_2)^2]^{3/2}} - x_2 + 3\cos t\,(x_2 \cos t + y_2 \sin t) \\ \ddot{y}_2 &= \frac{m_1(y_1-y_2)}{[(x_1-x_2)^2 + (y_1-y_2)^2]^{3/2}} - y_2 + 3\sin t\,(x_2 \cos t + y_2 \sin t)\ . \end{aligned}$$

From (5), we have in addition $m_1x_1 + m_2x_2 = 0$ and $m_1y_1 + m_2y_2 = 0$; so that (15) and (14) are not independent of each other. Setting $x = x_1 - x_2$ and $y = y_1 - y_2$, we find on subtraction of (15) from (17) that

$$(16)\qquad \begin{aligned} \ddot{x} &= \frac{-(m_1+m_2)x}{[x^2+y^2]^{3/2}} - x + 3\cos t\,(x\cos t + y \sin t) \\ \ddot{y} &= \frac{-(m_1+m_2)y}{[x^2+y^2]^{3/2}} - y + 3\sin t\,(x\cos t + y \sin t)\ . \end{aligned}$$

Setting $z = x + iy$, we can write (16) in the more compact form

$$\ddot{z} = \frac{-(m_1+m_2)z}{|z|^3} - z + 3\, e^{it}\, R(z\, e^{-it}) \, . \tag{17}$$

We introduce rotating coordinates at this stage by putting $\zeta = z\, e^{-it}$, so that $z = \zeta\, e^{it}$, $\dot{z} = (\dot{\zeta} + i\zeta)e^{it}$, $\ddot{z} = (\ddot{\zeta} + 2i\, \dot{\zeta} - \zeta)e^{it}$, and $|z| = |\zeta|$. Substituting in (17) and suppressing a factor e^{it}, we get

$$\ddot{\zeta} + 2i\, \dot{\zeta} - \zeta = \frac{-(m_1+m_2)\zeta}{|\zeta|^3} - \zeta + 3R(\zeta) \, .$$

Hence, setting $\zeta = \xi + i\,\eta$ and separating real and pure imaginary parts, we find that

$$\ddot{\xi} - 2\dot{\eta} = \frac{-(m_1 + m_2)\xi}{[\xi^2 + \eta^2]^{3/2}} + 3\xi$$

$$\ddot{\eta} + 2\dot{\xi} = \frac{-(m_1 + m_2)\eta}{[\xi^2 + \eta^2]^{3/2}}$$

which are Hill's equations. They appear in the more standard form

$$\begin{aligned} \ddot{x} - 2\dot{y} &= 3x - \frac{x}{(x^2+y^2)^{3/2}} \\ \ddot{y} + 2\dot{x} &= - \frac{y}{(x^2+y^2)^{3/2}} \end{aligned} \tag{18}$$

if we take $\xi = (m_1 + m_2)^{1/3}x$ and $\eta = (m_1 + m_2)^{1/3}y$. The present x and y are, of course, different from the x and y of (16).

I.3. Complete Decoupling

In carrying out the complete decoupling we observe that, if the k bodies are very close to each other compared to their distances from the other n-k bodies, their attraction (per unit of mass) on each other is much greater than on the other n-k bodies. This will cause greater relative accelerations of the k bodies than of the others. In order to exaggerate this effect we introduce the small parameter λ in such wise as to change the time scale as well as to multiply $r_1, \ldots, r_k$ and simultaneously to render those terms on the right of (12) which involve the other r's to be very small compared with the other terms. This is achieved by multiplying $r_1, \ldots, r_k$ by λ^2 and t by λ^3, and leaving $r_{k+1}, \ldots, r_n$ and the masses untouched. When (12) is thus modified, we find that

$$\lambda^{-4} \ddot{r}_i = \lambda^{-4} \sum_{j \neq i} \frac{m_j(r_j - r_i)}{|r_j - r_i|^3} + \lambda^2 \sum_{\alpha = k+1}^{n} \left(- \frac{m_\alpha r_i}{|r_\alpha|^3} + \frac{3m_\alpha r_\alpha (r_\alpha \cdot r_i)}{|r_\alpha|^5}\right) .$$

First multiplying by λ^4 and then letting $\lambda \to 0$ we get the complete decoupling with (12) replaced by

$$\ddot{r}_i = \sum_{j \neq i} \frac{m_j(r_j - r_i)}{|r_j - r_i|^3}, \qquad i = 1, \ldots, k. \tag{19}$$

which are precisely the equations for the k-body problem.

Notice that the same type of reduction occurs if we take the Hill equations (18) in rotatingcoördinates. If we have multiply x and y by λ^2 and t by λ^3, these equations (17) take the modified form (after multiplying through by λ^4)

$$\ddot{x} - 2\lambda^3\dot{y} = 3\lambda^6 x - \frac{x}{(x^2 + y^2)^{3/2}}$$

$$\ddot{y} + 2\lambda^3\dot{x} = - \frac{y}{(x^2 + y^2)^{3/2}}$$

which reduce in the limit $\lambda \to 0$ to the equations for Keplerian motion.

It does not matter in such approximations whether rotating or non-rotating coordinates are used. Correspondingly the resulting approximations can be expected to be valid only over time intervals that are short relative to the period of the rotation although they could be long relative to the period (say) of the Keplerian elliptic motions representing motions (in the Hill case) in which m_1 and m_2 are very close to each other.

All of this is not in the least surprising, since the same change in time scale applied to the equation (10) modified the latter merely by introducing a factor λ^{-6} on the left. Hence, on multiplying by λ^6 and then allowing λ to approach zero, we get

$$\ddot{r}_\alpha = 0\,, \qquad \alpha = k+1, \ldots, n \tag{20}$$

as the limiting form of the other decoupled system. The time interval for a valid approximation of (12) by (19) would presumably be comparable to the time interval for a valid approximation of (10) by (20).

I.4. Mathematical Interpretation

The above intuitive discussion, about the "introduction of parameters" and the "exaggeration" of certain effects by allowing the parameters to approach their limits, may have had the desired effect of minimizing the complication of formulas; but it has resulted in a logically inconsistent notation; and it may possibly have baffled the reader in other respects as well. In order to clarify this situation we offer the following mathematical interpretation of the main result of this Section.

Consider the n-body problem with the n masses denoted by $m_1, m_2, \ldots, m_k, s^3m_{k+1}, s^3m_{k+2}, s^3m_{k+3}, \ldots, s^3m_n$. Let the position vectors of these n bodies relative to the center of gravity of the first k of them be denoted respectively by $\lambda^2r_1, \ldots, \lambda^2r_k, sr_{k+1}, \ldots, sr_n$ and let the time be denoted by λ^3t. We regard s and λ as constant (scalar) parameters. In this rather unfamiliar notation the usual equations for the n-body problem (taking the gravitational constant to be unity) may be written out. Thus the r's, considered as functions of t, are found to satisfy the system,

$$\begin{aligned}\frac{d^2r_i}{dt^2} &= \sum_{j\neq i} \frac{m_j(r_j-r_i)}{|r_j-r_i|^3} \\ &+ \sum_{\alpha=k+1}^{n} \frac{\lambda^4s^3m_\alpha(sr_\alpha-\lambda^2r_i)}{|sr_\alpha-\lambda^2r_i|^3} - \mu^{-1}\sum_{j=1}^{k}\sum_{\alpha=k+1}^{n}\frac{\lambda^4s^3m_jm_\alpha(sr_\alpha-\lambda^2r_j)}{|sr_\alpha-\lambda^2r_j|^3}\end{aligned} \tag{21}$$

$$\frac{d^2r_\alpha}{dt^2} = \sum_{j=1}^{k} \frac{s^{-1}\lambda^6 m_j(\lambda^2 r_j - sr_\alpha)}{|\lambda^2 r_j - sr_\alpha|^3} + \sum_{\beta\neq\alpha} \frac{s^2\lambda^6 m_\beta(sr_\beta - sr_\alpha)}{|sr_\beta - sr_\alpha|^3}$$

(22)

$$- \lambda^6\mu^{-1} \sum_{j=1}^{k} \sum_{\beta=k+1}^{n} \frac{s^2 m_j m_\beta(sr_\beta - \lambda^2 r_j)}{|\lambda^2 r_j - sr_\beta|^3} ,$$

where $\mu = m_1 + m_2 + \ldots + m_k$. Letting $s \to \infty$, equations (21) take the form

$$\frac{d^2r_i}{dt^2} = \sum_{j\neq i} \frac{m_j(r_j - r_i)}{|r_j - r_i|^3} + \sum_{\alpha=k+1}^{n} \lambda^6\left(- \frac{m_\alpha r_i}{|r_\alpha|^3} + \frac{3m_\alpha r_\alpha(r_\alpha \cdot r_i)}{|r_\alpha|^5}\right) . \tag{23}$$

and equations (22) take the form

$$\frac{d^2r_\alpha}{dt^2} = \lambda^6 \left(\sum_{\beta\neq\alpha} \frac{m_\beta(r_\beta - r_\alpha)}{|r_\beta - r_\alpha|^3} - \sum_{\beta=k+1}^{n} \frac{m_\beta r_\beta}{|r_\beta|^3}\right) . \tag{24}$$

Equations (23) and (24) are, of course, mere modifications of equations (12) and (10) respectively. Namely they contain the parameter λ, since the time is here denoted by $\lambda^3 t$ instead of t as in (12) and (10). Partial decoupling has been achieved in the limit $s \to \infty$, since the equations (24) are independent of $r_1, \ldots, r_k$. But the equations (23) for $\lambda \neq 0$, still contain all n of the unknown vectors. Complete decoupling only occurs in the limit $\lambda \to 0$, leading to the systems displayed in formulas (19) and (20).

II. QUASI FIRST INTEGRALS

II.1. Definition

Consider the system

$$\dot{x} = f(\sigma, x, y), \qquad \dot{y} = g(\sigma, x, y) , \tag{1}$$

where x and f are N-vectors, where y and g are K-vectors, where σ is a scalar parameter, and where the dot denotes differentiation with respect to the independent variable t. We also assume that f and g are of class C' in the region where solutions are considered. The system (1) is evidently of order $N + K$.

Suppose that this system is partially decoupled when σ takes on a specific value, say 0, in the sense that $g(0, x, y)$ is independent of x. We write

$$g(0, x, y) = G(y) , \tag{2}$$

where G is a K-vector. Let $y(t)$ denote any fixed solution of the "decoupled" system

$$\dot{y} = G(y) \tag{3}$$

and consider the "non-homogeneous variational equations"

$$\dot{Y} = G_y[y(t)]Y + g_\sigma(0, x, y(t)) , \tag{4}$$

where Y is a K-vector, where G_y is the Jacobian matrix of the components of G with respect to the components of y, and where g_σ is the partial derivative of g with respect to σ.

<u>A function</u> $F(t, x, Y)$ <u>is said to be a quasi-first integral of the system</u>

$$\dot{x} = f(0, x, y(t)) \tag{5}$$

<u>of order</u> N, <u>if it is an ordinary first integral of the</u> $(N + K)$<u>th order system consisting of</u> (4) <u>and</u> (5).

In other words, if $x(t)$ is any solution of (5) and if the pair $[x(t), Y(t)]$ satisfies (4), then

$$F[t, x(t), Y(t)] \equiv \text{constant}. \tag{6}$$

Since (5) does not involve Y and since (4) is linear in Y, it is easy to express Y by quadratures in terms of x. In fact, if $\Omega(t)$ is a fundamental matrix solution of the homogeneous variational equations,

$$\dot{\Omega} = G_y[y(t)]\,\Omega ,$$

then any solution of (4) must assume the form

$$Y(t) = \Omega(t)\ [c + \int_0^t \Omega(s)^{-1}\, g_\sigma(0, x(s), y(s))\, ds] ,$$

where c is a suitably chosen constant K-vector. Conversely, if c is an arbitrary constant K-vector, the $Y(t)$ given by the last formula satisfies (4). Hence inserting this expression for $Y(t)$ into (6), we see that our definition of a quasi-first integral of (5) implies that

$$F[t, x(t), \Omega(t)(c + \int_0^t \Omega(s)^{-1} g_\sigma(0, x(s), y(s))ds)] \equiv \text{constant} ,$$

no matter how the constant K-vector c may be chosen.

We first wish to discuss this definition with respect to the system (21) and (22) of the previous section. These form a system of $3n$ equations, each of the second order, in the components of the r's. However, we have taken the origin at the center of gravity of $m_1, \ldots, m_k$, so that we have three known linear relationships among the unknown vectors $r_1, \ldots, r_k$. These relationships render three of the equations (21) redundant. Thus the system (21) and (22) provide only $3n - 3$ independent equations of the second order, $3k - 3$ from (21) and $3(n - k)$ from (22). These second order equations may each be replaced by two first order equations by the familiar devise of using phase space instead of configuration space. Equations (21) are therefore equivalent to a set of $N = 6k - 6$ first order equations and equations (22) are equivalent to a set of $K = 6(n - k)$ first order equations. It is with these understandings that we consider the system consisting of equations (21) and (22) of the preceding section as an example of the system (1) of the present section. We take $s = \sigma^{-1}$, so that as $\sigma \to 0$, the system becomes partially decoupled as already explained.

The full system of $N + K = 6n - 6$ equations admits four well known first integrals corresponding to the energy and the three components of angular momentum. (The other six integrals of the n body problem, namely those corresponding to the linear momentum, are not available because of the choice of our non-inertial coordinate system. They would be used for obtaining positions with respect to an inertial system in terms of unknown r's).

The quasi-first integrals were invented to investigate what happens to these first integrals in the partially decoupled system. Briefly as $\sigma \to 0$, all four of these first integrals reduce to first integrals of the system (24) of the previous section; while the system (23), considered for a fixed solution of (24) in which $r_\alpha(t)$ are regarded as known for $\alpha = k + 1, \ldots, n$, has, in general, no first integrals whatever. It does turn out, however, to possess four quasi-first integrals in a manner described in a more general setting below.

II. 2. Theorems on Quasi-First Integrals

In the sequel we suppose that the system (1) admits a first integral $h(\sigma, x, y)$ and that $h(0, x, y)$ is independent of x. Let

$$(7) \qquad h(0, x, y) = H(y) .$$

We also assume that both h and H are of class C''.

Theorem 1. If $y(t)$ is any fixed solution of (3), then the function

(8) $$F(t, x, Y) = h_\sigma(0, x, y(t)) + H_y(y(t))Y$$

is a quasi-first integral of the system (5).

Proof. Since $h(\sigma, x, y)$ is a first integral of (1), we have

$$h_x(\sigma, x, y)\, f(\sigma, x, y) + h_y(\sigma, x, y)\, g(\sigma, x, y) \equiv 0$$

Differentiate this with respect to σ and then set $\sigma = 0$. Remembering (2) and (7), we find in this way that

(9) $$h_{\sigma x}(0, x, y)\, f(0, x, y) + h_{\sigma y}(0, x, y)\, G(y) + H_y(y)\, g_\sigma(0, x, y) \equiv 0 ,$$

since $h_x(0, x, y) = 0$ by (7). Also the system (3) admits the function $H(y)$ as a first integral. Thus, we have the identity

$$H_y(y)\, G(y) \equiv 0 .$$

Differentiate this with respect to y and then set $y = y(t)$. We thus get

$$H_{yy}(y(t))\, G(y(t)) + H_y(y(t))\, G_y(y(t)) \equiv 0$$

Form the inner product of the left member of this last identity with the vector Y. We thus obtain

(10) $$Y\, H_{yy}(y(t))\, G(y(t)) + H_y(y(t))\, G_y(y(t))Y \equiv 0 ,$$

or more briefly

(11) $$[H_{yy}G + H_yG_y]Y \equiv 0 .$$

Now evidently, from (8) and (3), we have

(12) $$\frac{\partial F}{\partial t} = [\underline{h_{\sigma y}(0, x, y(t))} + H_{yy}(y(t))Y]\, G(y(t))$$

(13) $$\frac{\partial F}{\partial x}\, f(0, x, y(t)) = \underline{h_{\sigma x}(0, x, y(t))\, f(0, x, y(t))}$$

(14) $$\frac{\partial F}{\partial Y}\, \{G_y(y)Y + g_\sigma(0, x, y(t))\} = H_y(y(t))\, \{G_y(y)Y + \underline{g_\sigma(0, x, y)}\} .$$

We now add (12), (13), and (14). On the right hand side the underlined terms cancel out because of (9). The other terms cancel out because of (11) and the fact that $H_{yy}GY = H_{yy}YG$, the Hessian matrix H_{yy} being symmetric. Thus we find that

$$\frac{\partial F}{\partial t} + \frac{\partial F}{\partial x} f(0, x, y(t)) + \frac{\partial F}{\partial Y} \{G_y(y(t))Y + g_\sigma(0, x, y(t))\} \equiv 0 .$$

Hence F is a first integral of the system composed of (4) and (5), and hence, by the definition, it is a quasi-first integral of the system (5), as we set out to prove.

In the previous subsection we indicated how the Y can, in general, be eliminated from a quasi-first integral by use of quadratures acting on the unknown x. The method involved the knowledge of a fundamental matrix solution of the homogeneous variational equations. In the particular instance treated in Theorem 1, where Y occurs only in the combination $H_y(y(t))Y$ as in (8), this elimination is much easier. The essential facts are displayed as follows:

<u>Theorem 2.</u> If $y(t)$ is any fixed solution of the system

(15) $$\dot{y} = G(y) ,$$

which has the first integral $H(y)$, and if $Y(t)$ satisfies the non-homogeneous variational equations,

(16) $$\dot{Y} = G_y(y(t))Y + K(t) ,$$

then

(17) $$H_y(y(t))Y(t) = \int_{t_0}^{t} H_y(y(s))K(s)\,ds + H_y(y(t_0))Y(t_0) .$$

In particular $H_y(y(t))Y$ is a first integral of the homogeneous variational equations (cf. special case $K(t) \equiv 0$).

<u>Proof.</u> Since $H(y)$ is a first integral of (15), we know that $H_y(y)G(y) \equiv 0$ is an identity in y. Differentiating with respect to y and then setting $y = y(t)$, we obtain the following identity in t:

$$H_{yy}(y(t))\, G(y(t)) + H_y(y(t))\, G_y(y(t)) \equiv 0 .$$

Forming the inner product of the left side of this identity with the vector $Y(t)$ we obtain

$$H_{yy}(y(t))\; G(y(t))\; Y(t) + H_y(y(t))\; G_y(y(t))\; Y(t) \equiv 0\;. \tag{18}$$

On the other hand

$$\frac{d}{dt}\,[H_y(y(t))\; Y(t)] = H_{yy}(y(t))\dot{y}(t)Y(t) + H_y(y(t))\dot{Y}(t)\;.$$

Hence, from (15) and (16), we find that

$$\frac{d}{dt}\,[H_y(y(t))Y(t)] = H_{yy}(y(t))G(y(t))Y(t) + H_y(y(t))G_y(y(t))Y(t) + H_y(y(t))K(t)\,.$$

But the first two terms on the right of this equation cancel because of (18). What is left is equivalent to the stated theorem.

Theorem 3. Under the hypotheses of Theorem 1, let $x(t)$ be an arbitrary solution of (5). Then

$$h_\sigma(0,\, x(t),\, y(t)) + \int_{t_0}^{t} H_y(y(s))g_\sigma(0,\, x(s),\, y(s))ds = \text{constant.}$$

Proof. Since the F of formula (8) is a known quasi-first integral of (5), it is known at once that

$$h_\sigma(0,\, x(t),\, y(t)) + H_y(y(t))\; Y(t) = \text{constant}\;, \tag{19}$$

where $Y(t)$ is any vector satisfying (4), x being set equal to $x(t)$. But we then know from Theorem 2, with $K(t) = g_\sigma(0,\, x(t),\, y(t))$, that

$$H_y(y(t))Y(t) = \int_{t_0}^{t} H_g(y(s))\; g_\sigma(0,\, x(s),\, y(s))ds + \text{constant}\;.$$

Inserting into (19), we obtain the result to be proved.

II.3. Application to Reduced 3-Body Problem

The detailed application of the theory of quasi-first integrals to the systems of the previous Section has not been carried out. Instead we present the following outline of how the application may be made to the planar reduced three body problem and in particular to the restricted problem, at least to the point of producing, in this example, the constants of motion referred to in Theorem 3.

We consider the planar three body problem with three masses σ, μ, $1 - \mu$. We use two frames of reference, one with origin at the center of mass of the three bodies, and a parallel frame with origin at the body with mass μ. We let x_1, x_2 be the coordinates of σ with respect to the first frame, and we let y_1, y_2 be the coordinates of $1-\mu$ with respect to the second frame. Then it is elementary, though somewhat laborious, to show that the full planar three body problem may be reduced to the solution of the following system of differential equations:

$$\ddot{x}_i = -\mu[(\sigma+1)x_i + (1-\mu)y_i]R^{-3} - (1-\mu)[(\sigma+1)x_i - \mu\, y_i]r^{-3} \tag{20}$$
$$i = 1, 2.$$

$$\ddot{y}_i = -y_i[y_1^2 + y_2^2]^{-3/2} + \sigma[(\sigma+1)x_i - \mu y_i]r^{-3} - \sigma[(\sigma+1)x_i + (1-\mu)y_i]R^{-3} \tag{21}$$
$$i = 1, 2.$$

where, both in (20) and in (21), we have used the abbreviations,

$$\begin{aligned} r^2 &= [(\sigma+1)x_1 - \mu y_1]^2 + [(\sigma+1)x_2 - \mu y_2]^2 \\ R^2 &= [(\sigma+1)x_1 + (1-\mu)y_1]^2 + [(\sigma+1)x_2 + (1-\mu)y_2]^2 \end{aligned} \tag{22}$$

The four second order equations displayed in (20) and (21) are equivalent to eight first order equations, which form the system to be identified in this example with the System (1). Thus $N = K = 4$ and x and y are to be regarded as four vectors with components $(x_1, x_2, \dot{x}_1, \dot{x}_2)$ and $(y_1, y_2, \dot{y}_1, \dot{y}_2)$ respectively. The system has two first integrals,

$$h^{(1)} = \sigma(\sigma+1)(x_1\dot{x}_2 - \dot{x}_1x_2) + \mu(1-\mu)(y_1\dot{y}_2 - \dot{y}_1y_2)$$

and

$$h^{(2)} = \frac{1}{2}\sigma(\sigma+1)(\dot{x}_1^2 + \dot{x}_2^2) + \frac{1}{2}\mu(1-\mu)(\dot{y}_1^2 + \dot{y}_2^2) - \frac{\mu(1-\mu)}{\sqrt{y^2+y_2^2}} - \frac{\mu\sigma}{R} - \frac{(1-\mu)\sigma}{r}$$

Thus, we obtain

$$
(23) \qquad
\begin{aligned}
h_\sigma^{(1)}\,(0,\, x,\, y) &= x_1\dot{x}_2 - \dot{x}_1 x_2 \\
h_\sigma^{(2)}\,(0,\, x,\, y) &= \tfrac{1}{2}\,(\dot{x}_1^{\,2} + \dot{x}_2^{\,2}) - \frac{\mu}{R_0} - \frac{1-\mu}{r_0}
\end{aligned}
$$

where R_0 and r_0 are obtained from the expressions for R and r respectively by setting $\sigma = 0$ in (22).

$$
(24) \qquad
\begin{aligned}
H^{(1)}(y) &= h^{(1)}(0,\, x,\, y) = \mu(1-\mu)(y_1\dot{y}_2 - \dot{y}_1 y_2) \\
H^{(2)}(y) &= h^{(2)}(0,\, x,\, y) = \tfrac{1}{2}\,\mu(1-\mu)(\dot{y}_1^{\,2} + \dot{y}_2^{\,2}) - \frac{\mu(1-\mu)}{\sqrt{y_1^{\,2}+y_2^{\,2}}}
\end{aligned}
$$

The system (3) in this example is represented by (21) with $\sigma = 0$, namely

$$
(25) \qquad \ddot{y}_i = -\,y_i\,[y_1^{\,2} + y_2^{\,2}]^{-3/2}\,, \qquad i = i,\ 2.
$$

and the system (5) is represented by

$$
(26) \qquad \ddot{x}_i = -\,\mu[x_i + (1-\mu)y_i]R_0^{-3} - (1-\mu)[x_i - \mu y_i]r_0^{-3}\,, \qquad i = i,\ 2,
$$

these being the equations for the reduced problem of three bodies when y_1 and y_2 are thought of as being known functions of t satisfying (25). The four-vector $g\,(0,\, x,\, y)$ is readily seen, in this example, to have the components,

$$
0,\ 0,\ (x_1 - \mu y_1)r_0^{-3} - (x_1 + (1-\mu)y_1)R_0^{3},\ (x_2-\mu y_2)r_0^{-3} - (x_2+(1-\mu)y_2)R_0^{-3}\,.
$$

According to Theorem 3, we should have two constants of motion, which we shall denote by c_1 and c_2, given by the formula,

$$
c_i = h_\sigma^{(i)}(0,\, x(t),\, y(t)) + \int_{t_0}^{t} H_y^{(i)}(y(s))\; g_\sigma(0,\, x(s),\, y(s))\,ds, \qquad i = 1,\ 2.
$$

We now have available all the data necessary for the use of this formula, namely the four components of $g_\sigma(0,\, x,\, y)$, and the expressions for $h_\sigma^{(i)}(0,\, x,\, y)$

and $H_y^{(i)}(y)$, given by (23) and (24) respectively. We thus obtain

$$c_1 = x_1\dot{x}_2 - \dot{x}_1 x_2 + \mu(1-\mu)\int_{t_0}^{t} (x_2 y_1 - x_1 y_2)(r_0^{-3} - R_0^{-3})ds \tag{27}$$

$$\begin{aligned} c_2 = {} & \tfrac{1}{2}(\dot{x}_1^{\,2} + \dot{x}_2^{\,2}) - \frac{\mu}{R_0} - \frac{(1-\mu)}{r_0} + \mu(1-\mu)\int_{t_0}^{t} (x_1\dot{y}_1 + x_2\dot{y}_2)(r_0^{-3} - R_0^{-3})ds \\ & - \mu(1-\mu)\int_{t_0}^{t} (y_1\dot{y}_1 + y_2\dot{y}_2)(\mu\, r_0^{-3} + (1-\mu)R_0^{-3})ds \end{aligned} \tag{28}$$

where the symbols x_i, $\dot{x}_i$, y_i $\dot{y}_i$, r_0, R_0, occurring outside the integral signs, represent functions of the independent variable t; but, when they occur inside the integral signs, they represent the corresponding functions of s, the variable of integration.

In the special case of the restricted problem of three bodies, in which we take $y_1(t) = \cos t$ and $y_2(t) = \sin t$, we have $y_1\dot{y}_1 + y_2\dot{y}_2 = 0$, so that the second of the two integrals in (28) drops out completely. Moreover $\dot{y}_1 = -y_2$ and $\dot{y}_2 = y_1$. This means that the first of the two integrals in (28) is the same as the only integral in (27). Thus we find that

$$c_2 - c_1 = \frac{1}{2}(\dot{x}_1^{\,2} + \dot{x}_2^{\,2}) - (x_1\dot{x}_2 - \dot{x}_1 x_2) - \frac{\mu}{R_0} - \frac{(1-\mu)}{r_0} \tag{29}$$

$$c_1 = x_1\dot{x}_2 - \dot{x}_1 x_2 + \mu(1-\mu)\int_{t_0}^{t} (x_2(s)\cos s - x_1(s)\sin s)(r_0^{-3} - R_0^{-3})ds\ . \tag{30}$$

Thus in this particular case, we have a first integral in the ordinary sense, namely the integral of Jacobi. In rotating coördinates (ξ_1, ξ_2), connected with the non-rotating coördinates by the transformation,

$$(x_1 + i\, x_2) = (\xi_1 + i\, \xi_2)\, e^{it}\ ,$$

the equations (26), in which $y_1 = \cos t$ and $y_2 = \sin t$, become

$$\ddot{\xi}_1 - 2\dot{\xi}_2 = \xi_1 - \frac{\mu(\xi_1+1-\mu)}{R_0^3} - \frac{(1-\mu)(\xi_1-\mu)}{r_0^3}$$

(31)

$$\ddot{\xi}_2 + 2\dot{\xi}_1 = \xi_2 - \frac{\mu\xi_2}{R_0^3} - \frac{(1-\mu)\xi_2}{r_0^3} \quad ,$$

where R_0 and r_0 have the same meaning as before but are now expressed in terms of (ξ_1, ξ_2) instead of (x_1, x_2), so that

$$R_0^2 = (\xi_1+1-\mu)^2 + \xi_2^2 \quad \text{and} \quad r_0^2 = (\xi_1-\mu)^2 + \xi_2^2 \ .$$

These are the usual equations for the restricted problem of three bodies. The equation (29) is transformed into

$$\frac{1}{2}(\dot{\xi}_1^2+\dot{\xi}_2^2) - \frac{1}{2}(\xi_1^2+\xi_2^2) - \frac{\mu}{R_0} - \frac{(1-\mu)}{r_0} = c_2 - c_1 \ ,$$

which is the usual form of the integral of Jacobi. Finally the equation (30) is transformed into

$$\xi_1^2 + \xi_2^2 + \xi_1\dot{\xi}_2 - \dot{\xi}_1\xi_2 + \mu(1-\mu)\int_{t_0}^{t} \xi_2(s)\ [r_0^{-3} - R_0^{-3}]ds = c_1 \ .$$

The fact that the left member of this equation is a constant of the motion is an easy direct consequence of (31).

III. ERROR ESTIMATION

III.1. The Hyperbolic Cosine Estimate

We wish now to present a method for estimating the error in the computation of trajectories by using such decoupled systems, as those considered in Section I, instead of the exact systems. Actually the method applies to any kind of approximation, whether partial or complete decoupling occurs, or not. The problem is formulated as follows:

Suppose we have two N-vector functions $f(x)$ and $g(x)$ of the N-vector x, defined and continuous in some region R of N-vector space. Suppose also that f and g are approximations to each other in the sense that the norm of their difference is bounded throughout R by some positive number δ. The smaller δ is, the better the approximation. We thus assume that in R

(1) $$||f(x) - g(x)|| \leq \delta.$$

The norm $||\ldots||$ referred to may be chosen arbitrarily, so long as it has the usual properties of a norm in N-dimensional vector space.

Let $x(t) \in R$ and $y(t) \in R$ be solutions of the systems

(2) $$\frac{d^2x}{dt^2} = f(x)$$

and

(3) $$\frac{d^2y}{dt^2} = g(y)$$

assuming the same initial conditions

(4) $$x(0) = y(0) = a, \text{ say, and } \dot{x}(0) = \dot{y}(0) = b,$$

and defined on the interval $0 \leq t \leq T$. The problem is to find as small an upper bound as possible for $||x(t) - y(t)||$.

In the application to the n-body problem, $N = 3n$, and the region R is a region of configuration space in which the distance between each pair of the bodies exceeds a positive number assigned to such pair.

We could write the systems (2) and (3) as systems of first order equations by doubling the dimensionality of the space, and then we could read off from the classical literature estimates of the required type, at least, if one of the vector functions f or g satisfies a Lipschitz condition,

(5) $$||f(x') - f(x)|| \leq B\,||x' - x||.$$

We prefer, however, to adapt the classical methods directly to the systems of second order equations in order to get better results. (cf. E. Kamke, Differentialgleichungen, Lösungsmethoden, und Lösungen. Chelsea Publishing Company, New York, 1948, pp. 40-41.) We shall indeed assume that (5) holds as long as x and x' are both in R, and, for ease in later formulations, we introduce the following definition.

An N-vector function f defined in a region R of N-dimensional vector space and satisfying the Lipschitz condition (5) with Lipschitz constant B is said to satisfy the <u>Extension Hypothesis relative to R and B</u>, if for every positive number $A > B$ it is possible to define an N-vector function f_A throughout the whole of N-dimensional vector space such that

(6) $$f_A(x) \equiv f(x)$$

for all $x \in R$ and

$$(7) \qquad \|f_A(x') - f_A(x)\| \leq A \, \|x' - x\|$$

for any two N-vectors x and x', whether in R or not.

Theorem 1. If f satisfies the Extension Hypothesis, relative to R and B, and if (1), (2), (3), (4) are also assumed, then

$$(8) \qquad \|x(t) - y(t)\| \leq \delta B^{-1} (\cosh (B^{\frac{1}{2}} t) - 1) \qquad \text{for } 0 \leq t \leq T.$$

Proof. $x(t)$ and $y(t)$ satisfy respectively the systems of integral equations

$$(9) \qquad x(t) = a + bt + \int_0^t f(x(s))(t-s)ds$$

and

$$(10) \qquad y(t) = a + bt + \int_0^t g(y(s))(t-s)ds$$

which under the initial conditions (4) are equivalent to the differential systems (2) and (3). Let

$$(11) \qquad \begin{aligned} &x_0(t) = y(t) \qquad \text{for } 0 \leq t \leq T \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ &x_k(t) = a + bt + \int_0^t f_A(x_{k-1}(s))(t-s)ds, \qquad k = 1, 2, \ldots . \end{aligned}$$

We note that all these successive approximations $x_1(t)$, $x_2(t)$, ... must exist for $0 \leq t \leq T$ even though some or all of them do not stay within the region R. For the function f_A is defined and continuous over the whole N-vector space. It is easy to prove that

$$(12) \qquad \lim_{k \to \infty} x_k(t) = x(t) \quad \text{uniformly on } [0, T] .$$

In fact, from (11) and (7), we find that

$$\|x_{k+1}(t) - x_k(t)\| \leq A \int_0^t \|x_k(t) - x_{k-1}(t)\| \, (t-s)ds .$$

If we now let $\mu = \max ||x_1(t) - x_0(t)||$ over the interval $[0, T]$, we may now prove easily by induction that

$$||x_{k+1}(t) - x_k(t)|| \leq A^k \mu t^{2k}/(2k)! \qquad \text{for } t \in [0, T] .$$

It follows from the Weierstrass test that the sequence $\{x_k(t)\}$ converges uniformly to some vector function $x^*(t)$, and, since, from (7) ,

$$||f_A(x_k(t)) - f_A(x^*(t))|| \leq A ||x_k(t) - x^*(t)||,$$

it is also obvious that

$$\lim_{k\to\infty} f_A(x_k(t)) = f_A(x^*(t)) \qquad \text{uniformly on } [0, T] .$$

Hence passing to the limit as $k \to \infty$, we see from (11) that

$$x^*(t) = a + bt + \int_0^t f_A(x^*(s))(t-s)ds .$$

This means that $x^*(t)$ satisfies the system $\ddot{x} = f_A(x)$ and the initial conditions (4). But $x(t)$, because of (6), also satisfies the same conditions. Because of the uniqueness theorems covering such solutions, it follows that $x^*(t) \equiv x(t)$. Thus (12) has now been established.

From (10) and (11) we are enabled to write

$$y(t) - x_k(t) = \int_0^t g(y(s))(t-s)ds - \int_0^t f_A(x_{k-1}(s))(t-s)ds .$$

whence, using the fact that $y(s) \in R$ so that $f_A(y(s)) = f(y(s))$ by (6), we find that

$$y(t) - x_k(t) = \int_0^t (g(y(s)) - f(y(s)))(t-s)ds + \int_0^t (f_A(y(s)) - f_A(x_{k-1}(s)))(t-s)ds.$$

Using (1) and (7), we now find that

$$||y(t) - x_k(t)|| \leq (\delta/2)t^2 + A \int_0^t ||y(s) - x_{k-1}(s)|| (t-s)ds.$$

From this we prove routinely by induction that

$$||y(t) - x_k(t)|| \leq \frac{\delta}{A} \sum_{p=1}^{k} \frac{A^p t^{2p}}{(2p)!}$$

Hence allowing $k \to \infty$, we find, from (12), that

$$||y(t) - x(t)|| \leq \frac{\delta}{A} \sum_{p=1}^{\infty} \frac{A^p t^{2p}}{(2p)!} = \delta A^{-1} (\cosh(A^{\frac{1}{2}} t) - 1) .$$

Since this is true for all $A > B$, we find easily that (8) must hold as stated by the theorem.

Theorem 2. The inequality (8) can not be improved under the hypotheses of Theorem 1.

Proof. We take the example in which $N = 1$, $||x|| = |x|$, $f(x) = Bx$, $g(x) = Bx + \delta$, and R is the set of all real numbers. Thus if (2), (3), and (4) are to be satisfied, we find at once that the difference $y(t) - x(t) = w(t)$, say, must satisfy $\ddot{w} = Bw + \delta$ together with the initial conditions $w(0) = \dot{w}(0) = 0$. Integrating, we find that

$$y(t) - x(t) = w(t) = \delta B^{-1} (\cosh (B^{\frac{1}{2}} t) - 1).$$

so that it is possible in particular examples for the equality sign in (8) to hold.

Theorem 3. Under the hypotheses of Theorem 1, we also have

$$||\dot{x}(t) - \dot{y}(t)|| \leq \delta B^{\frac{1}{2}} \sinh(B^{\frac{1}{2}} t) . \tag{13}$$

Proof. Subtracting (10) from (9) and differentiating, we see that

$$\dot{x}(t) - \dot{y}(t) = \int_0^t (f(x(s)) - g(y(s)))ds$$

$$= \int_0^t (f(x(s)) - f(y(s)))ds + \int_0^t (f(y(s)) - g(y(s)))ds$$

Hence, from (5) and (1), we obtain

$$||\dot{x}(t) - \dot{y}(t)|| \leq B \int_0^t ||x(s) - y(s)||ds + \delta t .$$

We now use Theorem 1 to obtain

$$||\dot{x}(t) - \dot{y}(t)|| \leq \delta \int_0^t (\cosh (B^{\frac{1}{2}} s) - 1) ds + \delta t = \delta B^{-\frac{1}{2}} \sinh B^{\frac{1}{2}} t ,$$

as we wished to prove.

III.2. A Useful Lemma

In the application of the preceding subsection to the equations of motion for the n-body problem it is necessary, first, to choose a region R in which a suitably chosen Lipschitz condition will hold, and, secondly to verify the Extension Hypothesis, introduced in Theorem 1. For both of these purposes the following Lemma is useful.

Lemma 1. Let $\phi_k(r) = r^{-1}$ if $r \geq 1$; but, for $0 \leq r \leq 1$, let

$$\phi_k(r) = \frac{(6k+3)(2k+3)}{8(k+1)k} - \frac{(2k+1)(2k+3)}{8k(k-1)} r^2 + \frac{3(2k+3)}{8k(k-1)} r^{2k} - \frac{3(2k+1)}{8(k+1)k} r^{2k+2} \quad \text{where } k > 1.$$

Then these definitions agree at $r = 1$. Moreover $\phi_k \in C'''$ for $0 \leq r < \infty$ and

$$|\phi_k^{(p)}(r)| \leq p + 0_p(k), \quad \text{where} \lim_{k\to\infty} 0_p(k) = 0, \qquad p = 1, 2.$$

Proof. To show that $\phi_k \in C'''$ it is only necessary to compute the right handed and left handed derivatives of orders 1, 2, 3 at the one point $r = 1$ where ϕ_k obviously fails to be analytic. It will appears that each left handed derivative at this point is equal to the corresponding right handed derivative. Moreover these right and left derivatives are also the limits of derivatives taken on the right and left respectively. The details are elementary and are omitted. We record, however, that for $0 \leq r \leq 1$.

$$\phi'_k(r) = -\frac{(2k+1)(2k+3)}{4k(k-1)} r + \frac{3(2k+3)}{4(k-1)} r^{2k-1} - \frac{3(2k+1)}{4k} r^{2k+1}$$

$$\phi''_k(r) = -\frac{(2k+1)(2k+3)}{4k(k-1)} + \frac{3(2k+3)(2k-1)}{4(k-1)} r^{2k-2} - \frac{3(2k+1)^2}{4k} r^{2k}$$

$$\phi'''_k(r) = \frac{3(4k^2+4k-3)}{2} r^{2k-3} - \frac{3(4k^2+4k+1)}{2} r^{2k-1}$$

Since $\phi_k(r)$ and its derivatives are, for $r \geq 1$, simple monotonic functions of r which tend to 0 as $r \to \infty$, the problem of estimating the maximum moduli of the first two derivatives of $\phi_k(r)$ may be reduced to a study of what happens on the interval $[0, 1]$.

Since $\phi'''_k(1) = -6$, since $\phi'''_k(r)$ has a simple zero at

$r = r_k = \left(\frac{4k^2+4k-3}{4k^2+4k+1}\right)^{1/2}$, and since it vanishes nowhere else on the open interval $(0, 1)$, it follows that $\phi''_k(r)$ assumes its maximum at $r = r_k$. But a straight forward calcultion based on the formula for $\phi''_k(r)$ shows that

$$\phi''_k(r_k) = -\frac{(2k+1)(2k+3)}{4k(k-1)} + \left(\frac{12k^2+12k+3}{4k(k-1)}\right)\left(\frac{4k^2+4k-3}{4k^2+4k+1}\right)^k$$

From this, we obtain by routine methods that

$$\lim_{k\to\infty} \phi''_k(r_k) = 2$$

Since $\lim\limits_{k\to\infty} \phi''_k(0) = -1$ and $\phi''_k(r)$ is monotonic increasing on $[0, r_k]$ and monotonic decreasing on $[r_k, \infty)$, it follows that

$$|\phi''_k(r)| \leq 2 + 0_2(k), \qquad \lim_{k\to\infty} 0_2(k) = 0 .$$

which is one of the two desired estimates. To get the other estimate we write the above formula for $\phi'_k(r)$ in the form

$$\phi'_k(r) = r\, h_k(r) \tag{13}$$

where

$$h_k(r) = -\frac{(2k+1)(2k+3)}{4k(k-1)} + \frac{3(2k+3)}{4(k-1)} r^{2k-2} - \frac{3(2k+1)}{4k} r^{2k} .$$

It is readily found that $h'_k(r)$ is positive throughout the open interval $(0, 1)$. Hence

$$-\frac{(2k+1)(2k+3)}{4k(k-1)} = h_k(0) \leq h_k(r) \leq h_k(1) = \phi'_k(1) = -1$$

for all r on the closed interval $[0, 1]$. Hence multiplying by r and using (13) we obtain, for $r \in [0, 1]$, the following inequalities

$$-\frac{(2k+1)(2k+3)}{4k(k-1)} \leqq -\frac{(2k+1)(2k+3)}{4k(k-1)}\, r \leqq r h_k(r) = \phi'_k(r) \leqq -r \leqq 0.$$

It follows that $|\phi'_k(r)| \leqq \frac{(2k+1)(2k+3)}{4k(k-1)} = 1 + 0_2(k)$, which is the other required estimate.

III.3. Lipschitz Constants for Non-Convex Regions

If f is an N-vector function of class C' defined over a convex region R in the N-dimensional vector space of the vectors x and if the N^2 partial derivatives of the components of f with respect to the components of x are bounded in R, then one can always find a Lipschitz constant B, which is related in a simple way to bounds on certain expressions involving these N^2 partial derivatives, such that (5) holds for any two points x' and x in R. For instance if P is the least common upper bound for the absolute values of all N^2 of these partial derivatives, one may take $B = NP$, at least if $||x||$ denotes the "Euclidean" norm $(\sum_{i=1}^{n} x_i^2)^{\frac{1}{2}}$, or the "uniform" norm $\max_{1 \leqq i \leqq N} (|x_i|)$, or the nameless norm $\sum_{i=1}^{n} |x_i|$.

If the region R is not convex the situation is more complicated. One way of dealing with it is to choose a convex region R^* which contains R and seek a C' extension f^* of f over the whole of R^* in which the bounds on these derivatives over R^* exceed the bounds over R by certain increments which are regarded as permissible. These permissible increments in some cases, which we term <u>regular</u> (but by no means in all cases), may be taken arbitrarily small. From these bounds over R^*, we then obtain a Lipschitz constant B^* such that

$$||f^*(x') - f^*(x)|| \leqq B^* \, ||x' - x|| \tag{14}$$

as long as x and x' are in R^* and hence a fortiori as long as they are in R where $f^* = f$.

In the <u>regular case</u>, just defined, the B^* would exceed by an arbitrarily small amount the B obtained in the same way from the bounds of the partial derivatives over R but ignoring the non-convexity of R where $f^* = f$. In other words if ε is any positive number we could (in the regular case) choose f^* in such a way that the B^* in (14) is not greater than $B + \varepsilon$; so that, remembering that f^* is always the same as f in R, we obtain, for any two points x' and x in R, the following inequality,

$$||f(x') - f(x)|| \leq (B + \varepsilon) \; ||x' + x|| \; .$$

Hence, allowing ε to tend to zero, we obtain (5). Thus in the regular case, the Lipschitz constant may be regarded as related to the bounds for the partial derivatives exactly as they should be if R were convex.

III.4. Application to the n-Body Problem

In the problem of n bodies, the $N = 3n$ components of the vector x are denoted by $(x_1, y_1, z_1, x_2, y_2, z_2, \ldots, x_n, y_n, z_n)$, where x_i, y_i, and z_i are the coördinates of the mass point m_i. The $3n$ components of f may be written in the form

$$\frac{1}{m_i}\frac{\partial U}{\partial x_i}\,, \quad \frac{1}{m_i}\frac{\partial U}{\partial y_i}\,, \quad \frac{1}{m_i}\frac{\partial U}{\partial z_i}\,, \qquad i = 1, 2, \ldots, n$$

where

$$U = \sum_{i<j} \frac{m_i\, m_j}{r_{ij}} \quad \text{and} \quad r_{ij} = ((x_i-x_j)^2 + (y_i-y_j)^2 + (z_i-z_j)^2)^{1/2} \; .$$

A typical closed region R to be considered is one which excludes all points for which one or more of the r_{ij} are zero, say the set of all points for which $r_{ij} \geq \sigma_{ij}$; $i, j = 1, 2, \ldots, n$; $i \neq j$. Here $\sigma_{ij} = \sigma_{ji}$ are $(1/2)n(n-1)$ assigned positive numbers. This region is evidently not convex; but we nevertheless are in the situation of the regular case mentioned above. In fact we can take R^* to be the whole $3n$-dimensional space and obtain a C' extension of f over the whole of R^* by choosing a positive integer $k > 1$ and taking the $3n$ components of the extension f^* to be

$$\frac{1}{m_i}\frac{\partial V_k}{\partial x_i}\,, \quad \frac{1}{m_i}\frac{\partial V_k}{\partial y_i}\,, \quad \frac{1}{m_i}\frac{\partial V_k}{\partial z_i}\,, \qquad i = 1, 2, \ldots n.$$

where

$$V_k = \sum_{i<j} m_i\, m_j\, \sigma_{ij}^{-1}\, \phi_k(\sigma_{ij}^{-1}\, r_{ij}) \; .$$

Then from the lemma of SubSection III.2 one finds that bounds for the partial derivatives of the f^* over the whole of R^* exceed the bounds for corresponding derivatives of f over R by arbitrarily small increments provided that k is taken sufficiently large.

It is evident also that this same substitution of V_k for U suffices to establish the Extension Hypothesis of Theorem 1, Sub-Section III.1.

For simplicity, we have considered here only the unreduced form of the equations for the n-body problem; but similar considerations would apply equally well also to such other forms as the helio-centric equations, the barycentric chain forms, or the various reduced equations considered in Section I.

PERTURBATION OF SOLUTIONS OF DIFFERENTIAL EQUATIONS SATISFYING GENERAL BOUNDARY CONDITIONS

By D. C. Lewis
Control Research Associates
Baltimore, Maryland

NASA Contract NAS 12-93

PERTURBATION OF SOLUTIONS OF DIFFERENTIAL EQUATIONS SATISFYING GENERAL BOUNDARY CONDITIONS

by

Daniel C. Lewis, Jr.
Professor of Applied Mathematics
The Johns Hopkins University

SUMMARY

The perturbation problem for a system of non-linear differential equations under general linear (but not necessarily homogenous) two point boundary conditions is reduced to the problem of solving k "bifurcation" equations, where k is the degeneracy of the problem. In the non-degenerate case $k = 0$, the set of such equations is vacuous and the problem is automatically solved. If $k \neq 0$ and if there are k independent first integrals, a significant transformation of the bifurcation equations in terms of these first integrals is carried out. The largest section of the paper is concerned with the construction of generalized Green's matrices (and with kindred matters) for linear systems with boundary conditions of arbitrary degeneracy. The generalized Green's matrix for the variational equations is prerequisite to our treatment of the non-linear problem.

INTRODUCTION

We extend here some results previously obtained for periodic boundary conditions. Cf. Annals of Mathematics, volume 63 (1956) pp. 535-548. Our original purpose was to develop general methods for the perturbation of periodic solutions of the n-body problem satisfying certain added conditions of symmetry. For this purpose the problem frequently could be formulated in terms of two point linear homogenous boundary conditions. It then became evident that the same methods might work also for problems of entirely different nature. Hence we devised the formulation described in Section I, where the boundary conditions are linear but not necessarily homogeneous. It is probable that much could be done for non-linear boundary conditions, but our historical introduction to the problem was such that questions of this sort arose too late to be considered here. There is obviously much still to be done in this field. With this in view we have carried out the developments of Section II on linear systems to a rather more complete state than strictly necessary for the primary purposes of this paper alone.

I. FORMULATION OF THE PROBLEM

We are concerned with the problem of solving a system of differential equations under general linear (but not necessarily homogeneous) two point boundary conditions. More precisely, let the differential system be written in the form,

$$d\xi/dt = F(t, \xi, \mu) ,$$

where t is the (scalar) independent variable, ξ is an n-vector the components of which are the unknown functions, and μ is a scalar parameter. F is an n-vector function defined and of class C' for $0 \leq t \leq T$, as long as μ and ξ belong to suitable domains. We seek functions $\xi(t)$ satisfying the above system as well as the following system of boundary conditions:

$$B_0\, \xi(0) + B_T\, \xi(T) = p ,$$

where p is a given n-vector and both B_0 and B_T are given $n \times n$ matrices. This system of boundary conditions is equivalent to n scalar conditions, which, if they are independent, are presumably just right to determine the n constants of integration for our n^{th} order system of differential equations. To insure the independence of our boundary conditions, we assume that the rank of the $n \times 2n$ matrix (B_0, B_T) is equal to n.

The perturbation problem, to which we devote attention, assumes that for a particular value of μ, which without loss of generality we may evidently take to be 0, we have a known solution, say $\zeta(t)$. Our problem is to determine a solution for $\mu \neq 0$, say $\xi(t, \mu)$, such that $\lim_{\mu \to 0} \xi(t, \mu) = \zeta(t)$.

We obtain a more convenient formulation of our problem by introducing the following notations:

$$x(t) = \xi(t) - \zeta(t)$$

$$A(t) = F_\xi(t, \zeta(t), 0)$$

$$f(t, x, \mu) = F(t, \zeta(t) + x, \mu) - F(t, \zeta(t, 0) - A(t)x ,$$

where F_ξ denotes the jacobian matrix of the components of F with respect to the components of ξ. Since $\zeta(t)$ is regarded as known, so is the matrix $A(t)$. And evidently $f(t, x, \mu)$ is of class C' and vanishes together with its partial derivatives with respect to the components of the vector x, when $\mu = 0$ and $x = 0$.

Since $\zeta(t)$ is supposed to be a solution of the two point boundary problem when $\mu = 0$, we have

$$d\zeta(t)/dt = F(t, \zeta(t), 0)$$

and

$$B_0 \zeta(0) + B_T \zeta(T) = p .$$

It is now clear that ξ will be a solution of the original two point boundary problem if and only if x satisfies the system

$$dx/dt = A(t)x + f(t, x, \mu)$$

and the homogeneous linear boundary conditions

$$B_0 x(0) + B_T x(T) = 0 .$$

Our perturbation problem can now be formulated in terms of finding a solution $x = x(t, \mu)$ of the last written differential sytem and the last written boundary conditions such that $\lim_{\mu\to 0} x(t, \mu) = 0$.

It is relatively easy to prove, for $|\mu|$ sufficiently small, the existence of such solutions provided that the so-called variational system

$$dx/dt = A(t)x$$

has no solutions satisfying the above homogeneous boundary conditions other than the "trivial" solution $x \equiv 0$. This is the so-called non-degenerate case. In this case the solution may always be found by a convergent process of successive approximations.

Many of the interesting problems in celestial mechanics are not, however, of this non-degenerate type. The problem is said to have degeneracy k, if the variational system has k linearly independent solutions satisfying the above homogeneous boundary conditions (upon which all other such solutions are linearly dependent). If $k > 0$, the original problem may have no solution; but, if it has, it may be found by a method of successive approximations followed by a solution of a system of k so-called "bifurcation" equations. The way in which this is done is explained in Section III. In the non-degenerate case $k = 0$, there are, of course, no bifurcation equations to solve.

To prepare the ground for the treatment in Section III, we must present in the next Section a considerable theory of linear systems. The non-degenerate case is included because all considerations are valid when $k = 0$. In fact most of the difficulties completely collapse when $k = 0$, many statements and conditions becoming vacuous.

The linear theory of Section II has close relationship with previous work by W. T. Reid on generalized Green's matrices for compatible systems of differential equations. (American Journal of Mathematics, volume 53 (1931), pages 443-459). In this reference there are references to still earlier work by G. A. Bliss and others. It would probably be possible to derive all the results we need by citing various theorems presented by these earlier authors. It will be easier, however, to give independent proofs; and this will have the advantage of developing some additional facts, which we shall also need.

II. PRELIMINARIES ON LINEAR SYSTEMS

We wish to consider continuous n-vector functions $x(t) = [x_1(t), \ldots, x_n(t)]$ defined on the interval $[0, T]$ and satisfying n indenendent linear homogeneous end conditions of the form,

$$B_0\, x(0) + B_T\, x(T) = 0\,, \tag{1}$$

where B_0 and B_T are given constant $n \times n$ matrices.

A vector $x(t)$ which satisfies (1) will be called <u>admissible</u>. Otherwise it will be called <u>inadmissible</u>.

The condition (1) can also be written in the form

$$(B_0, B_T)\begin{pmatrix} x(0) \\ x(T) \end{pmatrix} = 0 \tag{2}$$

where (B_0, B_T) is, of course, the $n \times 2n$-matrix formed by the indicated juxtaposition of B_0 and B_T and likewise

$$\begin{pmatrix} x(0) \\ x(T) \end{pmatrix}$$

is a $2n \times 1$ matrix formed by the juxtaposition of (column) vectors $x(0)$ and $x(T)$. Since the n conditions given by (1) or (2) are to be linearly independent, it is seen from (2) that it is both necessary and sufficient to assume the rank of (B_0, B_T) to be n, as we henceforth do. This means that (B_0, B_T) contains a non-sigular $n \times n$ matrix C. The matrix (B_0, B_T) may then be modified by permuting its columns so that it appears in the form (C, D). Then evidently

$$(C, D)\begin{pmatrix} C^{-1} \\ 0 \end{pmatrix} = I$$

where I is the $n \times n$ identity matrix. We thus find that there exist $n \times n$-matrices U_0 and U_T, such that

$$(B_0, B_T)\begin{pmatrix} U_0 \\ U_T \end{pmatrix} = I . \tag{3}$$

In fact, we may obtain $\begin{pmatrix} U_0 \\ U_T \end{pmatrix}$ by performing on the rows of $\begin{pmatrix} C^{-1} \\ 0 \end{pmatrix}$ the same permutation as we apply to the columns of (C, D) in order to recover (B_0, B_T). We also introduce the matrix $U(t)$ defined by

$$U(t) = (1 - tT^{-1})U_0 + tT^{-1}\, U_T , \tag{4}$$

so that $U(0) = U_0$ and $U(T) = U_T$. The following lemma then follows at once from (3).

Lemma 1. If (B_0, B_T) is of rank n, there exists an $n \times n$-matrix $U(t)$ whose elements are linear functions of t such that

$$B_0U(0) + B_TU(T) = I . \tag{5}$$

Incidentally, although the columns of $U(t)$ are all continuous n-vector functions of t, they are all inadmissible in the sense of the above definition of admissibility.

Theorem 1. Consider the linear differential system,

$$dx/dt = A(t)x + f(t) , \tag{6}$$

where x and f are n-vectors and A is an $n \times n$-matrix, A and f are known continuous functions of t, defined for $0 \leq t \leq T$. Let $X(t)$ be any $n \times n$-matrix such that $dX/dt = A(t)X$ and $\det X(0) \neq 0$ (and hence also $\det X(t) \neq 0$ for any t on $[0, T]$). Let B_0 and B_T be given constant $n \times n$-matrices such that the rank of (B_0, B_T) is n. Let $n - k$ denote the rank of the $n \times n$-matrix $B_0X(0) + B_TX(T)$, so that k is the number of linearly independent solutions of the homogeneous system corresponding to (6) which are admissible in the sense (as defined above) that they satisfy the boundary condition $B_0x(0) + B_Tx(T) = 0$.

Then there exist (independently of f) a $k \times n$-matrix function $\Xi(s)$ and an $n \times n$-matrix $G(t, s)$, both continuous, except that $G(t, s)$ posseses a finite jump at $t = s$, having the following three properties:

I. The system (6) possesses an admissible solution (i. e., a solution $x(t)$ such that $B_0 x(0) + B_T x(T) = 0$), if, and only if,

$$\int_0^T \Xi(s)\, f(s) ds = 0 . \tag{7}$$

II. If (7) is satisfied, the vector function

$$x(t) = \int_0^T G(t, s)\, f(s) ds \tag{8}$$

is a solution of (6) and is, moreover, the only admissible solution orthogonal to every admissible solution of the corresponding homogeneous system.

III. Whether (7) is satisfied or not, $x(t)$, defined by (8) is admissible.

<u>Proof.</u> Since $n - k$ is the rank of $B_0 X(0) + B_T X(T)$, there is a $k \times n$-matrix $\mathfrak{A}$ and an $n \times k$-matrix $\mathfrak{B}$, both of rank k such that

$$\mathfrak{A}[B_0 X(0) + B_T X(T)] = 0 \tag{9}$$

and

$$[B_0 X(0) + B_T X(T)]\mathfrak{B} = 0. \tag{10}$$

Here, of course, $0 \leq k \leq n$. The general solution of (6), whether admissible or not, is well known to be of the form

$$x(t) = X(t)\beta + \int_0^t X(t)\, X(s)^{-1}\, f(s) ds , \tag{11}$$

where the constants of integration are the components of the n-vector β. The attempt to find admissible solutions of (6) leads to the equation $B_0 x(0) + B_T x(T) = 0$, or

$$[B_0 X(0) + B_T X(T)]\beta + \int_0^T B_T X(T) X(s)^{-1} f(s) ds = 0 , \tag{12}$$

for the determination of β. On account of (9), this system of linear equations for the determination of β is consistent if, and only if,

$$\int_0^T \mathfrak{A} B_T X(T) X(s)^{-1} f(s) ds = 0 , \tag{13}$$

which can be written in the form (7), if we let

$$\Xi(s) = \mathfrak{A} B_T X(T) X(s)^{-1} . \tag{14}$$

If (13), or (7), is satisfied, it is possible to choose β in infinitely many ways so as to satisfy (12). Moreover, since the second term in (12) is the integral of a known matrix function of s multiplied into the vector $f(s)$, we see that β may be chosen to be given by a similar expression; say

$$\beta = \int_0^T P(s) f(s) ds$$

where $P(s)$ is a suitably chosen (not unique) continuous $n \times n$-matrix functin of s, independent of f. Substituting in (11) we have

$$x(t) = \int_0^T K(t, s) f(s) ds , \tag{15}$$

where $K(t, s) = X(t)[P(s) + X(s)^{-1}]$ for $0 \leq s < t$ and $K(t, s) = X(t)P(s)$ for $t \leq s \leq T$. The $x(t)$ given by (15) always satisfies (6) but is admissible if, and only if, (7) is satisfied.

We now wish to find another kernel matrix $\overline{K}(t, s)$, such that, upon writing

$$\overline{x}(t) = \int_0^T \overline{K}(t, s) f(s) ds , \tag{16}$$

we can say that the $\overline{x}(t)$ given by (16) is always admissible and satisfies (6) if (7) is satisfied. Since $x(t)$, as given by (15) is admissible if, and only if, (7) is satisfied, it is clear that (7) is equivalent to the condition

$$\int_0^T [B_0 K(0, s) + B_T K(T, s)] f(s) ds = 0 . \tag{17}$$

Hence, if we set

$$\overline{K}(t, s) = - U(t)[B_0 K(0, s) + B_T K(T, s)] + K(t, s) ,$$

so that

$$(17^*) \qquad \bar{x}(t) = -U(t)\int_0^T [B_0K(0, s) + B_TK(T, s)]\, f(s)ds + x(t) ,$$

then $\bar{x}(t)$ coincides with $x(t)$ whenever (17), or its equivalent (7), is satisfied. This is true for any $n \times n$-matrix function $U(t)$, but the following statement depends upon choosing $U(t)$ in accordance with Lemma 1. From (17*), we have

$$B_0\bar{x}(0) + B_T\bar{x}(T) = -[B_0U(0) + B_TU(T)]\int_0^T [B_0K(0, s) + B_TK(T, s)]f(s)ds$$

$$+ B_0x(0) + B_Tx(T) ,$$

which, in virtue of (15) and the fact that $[B_0U(0) + B_TU(T)] = I$ by Lemma 1, must vanish. Thus $\bar{x}(t)$, as defined by (16), must always be admissible, whether (7) is satisfied or not, as we wished to prove. And, of course, $\bar{x}(t)$ also satisfies (6) if (7) is fulfilled, because we already know that if (7) is fulfilled $\bar{x}(t)$ coincides with $x(t)$, which always satisfies (6).

Next we wish to find a kernel matrix $G(t, s)$ such that the $x(t)$ given by (8) satisfies the same properties as have already been specified for $\bar{x}(t)$ as given by (16) plus the additional property that $x(t)$ is always to be orthogonal to every admissible solution of $dx/dt = A(t)x$.

Assuming that (7) is satisfied, every admissible solution of (6) can evidently be written in the form

$$(18) \qquad x(t) = X(t)\mathfrak{B}\,\bar{c} + \int_0^T \bar{K}(t, s)f(s)ds ,$$

where $\bar{c}$ is a suitably chosen k-vector. This follows from the fact that the columns of $X(t)\mathfrak{B}$ form a complete set of admissible solutions of the homogeneous equations because of (10) and the fact that the rank of $B_0X(0) + B_TX(T)$ is $n - k$. Even when (7) is not satisfied, both terms on the right of (18) are known to be admissible and hence $x(t)$ as given by (18) is always admissible. We now show that it is always possible to choose $\bar{c}$ so that

$$(19) \qquad \int_0^T (X(t)\mathfrak{B})'\, x(t)dt = 0 ,$$

in other words, so that the $x(t)$ given by (18) is orthogonal to all the admissible solutions of the homogeneous equations and that this is true regardless of whether (7) holds. From (18) and (19) we see that it is necessary and sufficient for $\bar{c}$ to satisfy

$$\int_0^T \mathfrak{B}' X(t)'X(t)\mathfrak{B}\bar{c}\, dt + \int_0^T \mathfrak{B}' X(t)' \int_0^T \bar{K}(t, s)f(s)ds\, dt = 0$$

or

$$(20)\{\int_0^T \mathfrak{B}' X(t)'X(t)\mathfrak{B}\, dt\}\, \bar{c} = -\int_0^T \mathfrak{B}'(\int_0^T X(t)'\bar{K}(t, s)dt)f(s)ds \ .$$

We prove that the $k \times k$ matrix within the brackets { } is non-singular by showing in the following way that it is positive definite. Let $\bar{q}$ be any k-vector. Since $X(t)^{-1}$ exists, $\mathfrak{B}\bar{q} = 0$ if, and only if, $X(t)\mathfrak{B}\bar{q} = 0$. Since $\mathfrak{B}$ is of rank k, this is possible only if $\bar{q} = 0$. The quadratic form $\bar{q}(X(t)\mathfrak{B})'(X(t)\mathfrak{B})\bar{q}$ in the components of $\bar{q}$ is the sum of the squares of the components of $X(t)\mathfrak{B}\bar{q}$ and is, therefore, positive definite (in the real field). Hence

$$\int_0^T \bar{q}\mathfrak{B}'X(t)'X(t)\mathfrak{B}\bar{q}\, dt = \bar{q}\, (\int_0^T \mathfrak{B}'X(t)'X(t)\mathfrak{B}\, dt)\, \bar{q}$$

is also positive definite, considered as a quadratic form in the components of $\bar{q}$. Thus we now know that (20) can be solved for $\bar{c}$ in the form

$$(21) \qquad \bar{c} = \int_0^T \mathfrak{L}(s)f(s)ds \ ,$$

where $\mathfrak{L}(s)$ is (for a given $\bar{K}(t, s)$) a uniquely determined continuous $k \times n$-matrix function of s, independent of f. Substituting this value of $\bar{c}$ in (18), we finally arrive at (8) with

$$(22) \qquad G(t, s) = X(t)\mathfrak{B}\mathfrak{L}(s) + \bar{K}(t, s) \ .$$

The $G(t, s)$ just introduced by (22) thus satisfies all the requirements of the theorem; so that the proof of the latter is now complete.

<u>Lemma 2</u>. The rank of the $k \times n$-matrix $\mathfrak{A}B_T$ is k.

<u>Proof</u>. Suppose $\bar{q}$ is a k-vector such that

$$(23) \qquad \bar{q}\, \mathfrak{A}B_T = 0 \ .$$

It is enough to show that $\bar{q}$ must be the null vector, for this would mean that the k rows of B_T would be linearly independent, thus implying that its rank would be k. Multiplying on the right by $X(T)$ we have $\bar{q}\mathfrak{A}B_TX(T) = 0$. And, since by (9) it is known that $\mathfrak{A}B_TX(T) = -\mathfrak{A}B_0X(0)$, we also have $\bar{q}\mathfrak{A}B_0X(0) = 0$. Multiplying on the right by $X(0)^{-1}$, we find that

$$\bar{q}\,\mathfrak{A}B_0 = 0 . \tag{24}$$

Multiplying (24) and (23) respectively on the right by $U(0)$ and $U(T)$, and then adding, we obtain

$$\bar{q}\,\mathfrak{A}[B_0U(0) + B_TU(T)] = 0$$

But, since, by Lemma 1, the $n \times n$-matrix in the square brackets is just the identity matrix, we see at once that $\bar{q}\,\mathfrak{A} = 0$; and, since the rank of the $k \times n$-matrix $\mathfrak{A}$ is already known to be k, it follows of necessity that $\bar{q} = 0$, as we wished to prove.

<u>Theorem 2.</u> The rank of $\Xi(s) = \mathfrak{A}B_TX(T)X(s)^{-1}$ for each value of s on the interval $[0, T]$ is k.

<u>Proof.</u> This is a mere corollary of Lemma 2, since the rank of $X(T)$ is n and the rank of $X(s)^{-1}$ is also n.

It should be noted that neither Ξ nor G are uniquely determined by the requirements of Theorem 1. However we now prove the following

<u>Theorem 3.</u> Let $H(s)$ be any continuous $n \times k$-matrix function such that

$$C = \int_0^T \Xi(s)\, H(s)ds \tag{25}$$

is a non-singular $k \times k$-matrix. Then the $G(t, s)$ of Theorem 1 may be determined uniquely in such a manner that

$$\int_0^T G(t, s)\, H(s)ds = 0 \tag{26}$$

<u>Remark.</u> A possible choice for $H(s)$ is $\Xi(s)'$. For, in this case, we have

$$C = \int_0^T \Xi(s)\, \Xi(s)'ds ,$$

which, since $\Xi(s)$ is of rank k by Theorem 2, is easily proved to be a positive definite matrix and therefore non-singular.

Proof of Theorem 3. Let $G_0(t, s)$ be any kernel satisfying the requirements of Theorem 1, and let $R(t)$ be any $n \times k$-continuous matrix function whose columns are admissible. Thus, if we define

$$G(t, s) = G_0(t, s) - R(t) \quad \Xi(s) ,$$

It is seen that $G(t, s)$ also satisfies all the requirements of Theorem 1. If we choose

$$R(t) = (\int_0^T G_0(t, s) H(s)ds) C^{-1} ,$$

it is easy to verify that (26) is satisfied. Moreover, this $R(t)$ is admissible because $G_0(t, s)$ satisfies Property III of Theorem 1.

Let $\phi(s)$ be an arbitrary continuous vector function, and then set

$$f(s) = \phi(s) - H(s) C^{-1} \int_0^T \Xi(z)\phi(z)dz .$$

We easily see that this f satisfies (7), because of (25). Hence, from (26), we prove that

$$x(t) = \int_0^T G(t, s) f(s)ds = \int_0^T G(t, s) \phi(s)ds .$$

If $G^*(t, s)$ were a second matrix kernel satisfying the requirements of Theorem 1 as well as (26), we could also write

$$x^*(t) = \int_0^T G^*(t, s) f(s)ds = \int_0^T G^*(t, s) \phi(s)ds .$$

But $x^*(t) \equiv x(t)$ because of the uniqueness statement under Property II. Hence

$$\int_0^T [G(t, s) - G^*(t, s)] \phi(s)ds = 0$$

for arbitrary continuous ϕ. It follows that $G(t, s) \equiv G^*(t, s)$, thus completing the proof of Theorem 3.

Lemma 3. If B_0 and B_T are $n \times n$-matrices and if (B_0, B_T) is of rank n,

.hen there exist $n \times n$-matrices F_0 and F_T such that

$$B_0F_0 - B_TF_T = 0 \tag{27}$$

and such that $\begin{pmatrix} F_0 \\ F_T \end{pmatrix}$ is of rank n. Moreover F_0 and F_T are uniquely determined by these conditions up to a multiplication on the right by a common non-singular matrix.

Proof. Since (B_0, B_T) is of rank n, a suitable permutation on its columns yields an $n \times 2n$-matrix (C, D), where C is non singular. Then $(C, D)\begin{pmatrix} C^{-1} D G \\ - G \end{pmatrix} = DG - DG = 0$ regardless of the choice of the $n \times n$-matrix G. Construct $\begin{pmatrix} F_0 \\ -F_T \end{pmatrix}$ by performing the same permutation on the rows of $\begin{pmatrix} C^{-1} D G \\ -G \end{pmatrix}$ as that which when applied to the columns of (C, D) recovers (B_0, B_T). We thus have

$$B_0F_0 - B_TF_T = (B_0, B_T)\begin{pmatrix} F_0 \\ -F_T \end{pmatrix} = (C, D)\begin{pmatrix} C^{-1} D G \\ - G \end{pmatrix} = 0 ,$$

which establishes (27) for the constructed F's. Moreover rank of $\begin{pmatrix} F_0 \\ F_T \end{pmatrix} =$ rank of $\begin{pmatrix} F_0 \\ -F_T \end{pmatrix}$ = rank of $\begin{pmatrix} C^{-1} D G \\ - G \end{pmatrix}$, which is surely n, if, and only if, G is chosen to be non-singular. To see this, notice that any homogeneous linear relationship between the columns of G holds also for the columns of $C^{-1} D G$.

Suppose now we had two pairs F_0, F_T, and $\overline{F}_0$, $\overline{F}_T$ such that $B_0F_0 - B_TF_T = 0$, $B_0\overline{F}_0 - B_T\overline{F}_T = 0$, and with the ranks of both $\begin{pmatrix} F_0 \\ F_T \end{pmatrix}$ and $\begin{pmatrix} \overline{F}_0 \\ \overline{F}_T \end{pmatrix}$ equal to n. Then the same permutation applied to the columns of (B_0, B_T)

and to the rows of $\begin{pmatrix} F_0 \\ F_T \end{pmatrix}$ and $\begin{pmatrix} \overline{F}_0 \\ \overline{F}_T \end{pmatrix}$ sends the equations

$(B_0, B_T)\begin{pmatrix} F_0 \\ -F_T \end{pmatrix} = 0$ and $(B_0, B_T)\begin{pmatrix} \overline{F}_0 \\ -\overline{F}_T \end{pmatrix} = 0$ into the equations $(C, D)\begin{pmatrix} E \\ -G \end{pmatrix} = 0$

and $(C, D)\begin{pmatrix} \overline{E} \\ -\overline{G} \end{pmatrix} = 0$. These can be satisfied only by taking $E = C^{-1}DG$ and

$\overline{E} = C^{-1}D\overline{G}$. Since both $\begin{pmatrix} F_0 \\ F_T \end{pmatrix}$ and $\begin{pmatrix} \overline{F}_0 \\ \overline{F}_T \end{pmatrix}$ have rank n, we see as above that

both G and $\overline{G}$ must be non-singular. Hence $\overline{E} = EK$, where $K = G^{-1}\,\overline{G}$. Since also $\overline{G} = GK$, we see that

$$\begin{pmatrix} \overline{E} \\ -\overline{G} \end{pmatrix} = \begin{pmatrix} E \\ -G \end{pmatrix} K \quad .$$

Permuting the rows of $\begin{pmatrix} \overline{E} \\ -G \end{pmatrix}$ and $\begin{pmatrix} E \\ -G \end{pmatrix}$ by the inverse of the permutation last

mentioned leads to the result that

$$\begin{pmatrix} \overline{F}_0 \\ -\overline{F}_T \end{pmatrix} = \begin{pmatrix} F_0 \\ -F_T \end{pmatrix} K \quad ,$$

where $K = G^{-1}\,\overline{G}$ is non-singular. This establishes the last statement of the lemma.

A continuous vector function $x(t)$ defined on $[0, T]$ will be said to be <u>adjointly admissible</u> if $x(0)F_0 + x(T)F_T = 0$, where the F's satisfy the conditions expressed in Lemma 3. As a consequence of this lemma it makes no difference as to which particular F's are chosen. A change in the F's only amounts to replacing the last written equality by the equivalent relation $[x(0)F_0 + x(T)F_T]K = 0$, where K is non-singular.

<u>Theorem 4.</u> The rows of $\Xi(t)$ are adjointly admissible. In other words,

$$\Xi(0)F_0 + \Xi(T)F_T = 0 \quad . \tag{28}$$

<u>Proof.</u> We note from (9) that $\mathfrak{A}B_TX(T) = -\mathfrak{A}B_0X(0)$. Hence, from (14),

$\Xi(0) = \mathfrak{A}B_T X(T)X(0)^{-1} = -\mathfrak{A}B_0 X(0)X(0)^{-1} = -\mathfrak{A}B_0$, so that

$$\Xi(0) = -\mathfrak{A}B_0 . \tag{29}$$

Also, form (14), it is obvious that

$$\Xi(T) = +\mathfrak{A}B_T . \tag{30}$$

Hence, from (29) and (30), we see that $\Xi(0)F_0 + \Xi(T)F_T = -\mathfrak{A}(B_0F_0 - B_TF_T)$, which, by Lemma 3, must vanish. This establishes (28) and finishes the proof.

Theorem 5. The rows of $\Xi(t)$ constitute a complete set of linearly independent adjointly admissible solutions of the homogeneous linear system

$$d\xi/dt = -\xi A(t) \tag{31}$$

which is adjoint to the system $dx/dt = A(t)x$.

Proof. It is well known that the rows of $X(t)^{-1}$ satisfy (31). Since the rows of $\Xi(t) = \mathfrak{A}B_T X(T)X(t)^{-1}$, cf. (14), are merely linear combinations of the rows of $X(t)^{-1}$, the rows of $\Xi(t)$ must also satisfy (31). In Theorem 2, it was proved that the k rows of $\Xi(t)$ are linearly independent; and, in Theorem 4, it was proved that the rows are adjointly admissible. Hence Theorem 5 will be completely proved as soon as it is shown that k is the maximum number of linearly independent adjointly admissible solutions of (31). This is most easily done by noting that the relation between a system and its adjoint is a reciprocal one. The adjoint of the adjoint of a system is obviously the original system. The same may be said about admissibility and adjoint admissibility. A vector function which is adjointly admissible relative to the adjoint boundary conditions is admissible in the original sense. Hence, if there were k' ($> k$) linearly independent adjointly admissible solutions of the adjoint system, Theorems 1-4, applied to the adjoint system with adjoint boundary conditions, would tell us that there would have to be at least k' linearly independent admissible solutions of the original system $dx/dt = A(t)x$. But we already know that there are just k of them. This completes the proof.

III. BIFURCATION EQUATIONS

Our method of solving the problem formulated in Section I involves the preliminary solution of the system of integral equations

$$x(t, c, \mu) = X(t)\mathfrak{B}c + \int_0^T G(t, s)f[x(s, c, \mu), s, \mu]ds \; . \tag{32}$$

for the unknown function x. The solution will depend not only on t but also on the parameter μ and the k-vector c. Since G is bounded and the partial derivatives of f with respect to the components of x are continuous and vanish when x and μ do, it is almost self-evident that the following system of successive approximations

$$x_0(t, c, \mu) = X(t)\mathfrak{B}c$$

$$\cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot$$

$$x_m(t, c, \mu) = X(t)\mathfrak{B}c + \int_0^T G(t, s)f[x_{m-1}(s, c, \mu), s, \mu]ds, \qquad m \geq 1 \; .$$

$$\cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot$$

must, if $|\mu|$ and $||c||$ are sufficiently small, converge uniformly to a solution $x(t, c, \mu)$ such that $x(t, 0, 0) = 0$. Further details of the proof are indicated in a previous paper, The Role of First Integrals in the Perturbation of Periodic Solutions, Annals of Mathematics, volume 63 (1956), pp. 535-548, especially p. 545. Although in this previous paper only periodic boundary conditions were considered, the proof in the present more general case (in so far as it concerns the existence of solutions) is exactly the same. The main feature is that f must be Lipschitzian with arbitrarily small Lipschitz constant, if $|\mu|$ and $||x||$ are sufficiently restricted, because of the above mentioned properties of the partial derivatives of f. Assuming then that equation (32) is solved and that the solution is continuous (as it would have to be because of the uniform convergence of the continuous approximations), we are in a position to proceed to the next theorem.

Theorem 6. Let $x(t, c, \mu)$, where c is a k-vector, be a continuous solution of (32) such that $x(t, 0, 0) \equiv 0$. Then

$$B_0 x(0, c, \mu) + B_T x(T, c, \mu) = 0 \tag{33}$$

and $x(t, c, \mu)$ will satisfy

$$dx/dt = A(t)x + f(x, t, \mu) - H(t)\alpha \; , \tag{34}$$

where the relationship between H and G is as in Theorem 3 and where the k-vector α is given in the following formula:

$$\alpha = \alpha(c, \mu) = C^{-1} \int_0^T \Xi(t)f[x(t, c, \mu), t, \mu]dt \; . \tag{35}$$

<u>Proof.</u> Let $f(t) = f[x(t, c, \mu), t, \mu] - H(t)\alpha(c, \mu)$

$$\therefore \int_0^T \Xi(s)f(s)ds = \int_0^T \Xi(s)f[x(s, c, \mu), s, \mu]ds$$

$$- (\int_0^T \Xi(s)H(s)ds)C^{-1} \int_0^T \Xi(t)f[x(t, c, \mu), t, \mu]dt$$

$$= 0 \quad \text{since} \quad C = \int_0^T \Xi(s)H(s)ds \quad \text{as in Theorem 3.}$$

Hence the condition (7) is satisfied by the above defined $f(t)$. Hence the second term on the right of (32) satisfies (34) and is also admissible. Thus (33) is also satisfied. The first term on the right of (32) satisfies $dx/dt = A(t)x$ and because of (10) also satisfies (33). It is evident then that the sum $x(t, c, \mu)$ must have the stated properties.

It is clear from the theorem just proved that, if $c = c(\mu)$ can be chosen as a function of μ in such a way that

$$\alpha(c, \mu) = 0 , \tag{36}$$

then the differential system satisfied by $x(t, c(\mu), \mu)$ is the system $dx/dt = A(t)x + f(t, x, \mu)$ obtained from (34) and (36). Since the given two point boundary condition is also satisfied because of (33), it is seen that this $x(t, c(\mu), \mu)$ solves the problem formulated in Section I.

Thus the problem has been reduced to the problem of solving the system (36) of so-called bifurcation equations. This is generally very difficult. But, in case the differential equations admit some first integrals the bifurcation equations can be put into another form. We investigate this procedure when the number of first integrals is just equal to k.

Suppose that $\phi[x, t, \mu]$ is a k-vector, each component of which is a first integral of the system $dx/dt = A(t)x + f(x, t, \mu)$. Therefore

$$\phi_x(x, t, c, \mu)[A(t)x + f(x, t, \mu)] + \phi_t(x, t, \mu) \equiv 0.$$

Letting $x = x(t, c, \mu)$ and then integrating we get

$$\int_0^T \phi_x(x(t, c, \mu), t, c, \mu)[A(t)x(t, c, \mu) + f(x(t, c, \mu), t, \mu)] dt +$$
$$\int_0^T \phi_t(x(t, c, \mu), t, \mu)dt = 0.$$

On the other hand we have by the fundamental theorem of calculus

$$\phi[x(T, c, \mu), T, \mu] - \phi[x(0, c, \mu), 0, \mu] =$$

$$\int_0^T \phi_x(x(t, c, \mu)\dot{x}(t, c, \mu)dt + \int_0^T \phi_t(x(t, c, \mu), t, \mu)dt.$$

Hence, by subtraction, we find that

$$\phi[x(T, c, \mu), T, \mu] - \phi[x(0, c, \mu), 0, \mu] =$$

$$= \int_0^T \phi_x(x(t, c, \mu), t, \mu)[\dot{x}(t, c, \mu) - A(t)x(t, c, \mu) - f(x(t,c,\mu),t,\mu)]dt$$

$$= - \int_0^T \phi_x(x(t, c, \mu), t, \mu) H(t) \alpha(c,\mu) dt \quad \text{by (34)} .$$

We suppose that the components of $\phi(x, t, \mu)$ are <u>independent</u> first integrals at least for the solution $x(t, 0, 0) = 0$. This amounts to assuming that the $k \times n$-matrix $\phi_x(0, t, 0)$ has rank k. It is natural therefore to dispose of the arbitrariness of the $n \times k$-matrix $H(t)$, introduced first in Theorem 3 and used again in Theorem 6, by setting its transpose equal to $\phi_x(0, t, 0)$, at least provided that

$$C = \int_0^T \Xi(s)H(s)ds$$

turns out, for this choice of H, to be non-singular. Since therefore $\phi_x(0, t, 0) = H(t)'$ and since

$$\int_0^T H(t)' H(t)dt$$

is obviously positive definite so that it is, a fortiori, non-singular, it is seen by continuity that the matrix

$$\int_0^T \phi_x(x(t, c, \mu), t, \mu)H(t)dt$$

must also be non-singular for sufficiently small $||c||$ and $|\mu|$.

It now follows from the last equality that the so-called "bifurcation equation", $\alpha(c, \mu) = 0$ is equivalent to the equation

$$\phi[x(T, c, \mu), T, \mu] - \phi[x(0, c, \mu), 0, \mu] = 0 ,$$

at least if $||c||$ and $|\mu|$ are sufficiently small.

In the periodic case where $\phi(x, t, \mu) \equiv \phi(x, t + T, \mu)$ and $x(0, c, \mu) = x(T, c, \mu)$, the above result shows that the bifurcation equations are identically satisfied. But this happy circumstance apparently need not occur in general.

HIGH ORDER TIME DERIVATIVES OF POWERS OF THE RADIUS VECTOR

By P. Sconzo and D. Valenzuela
IBM Center for Exploratory Studies
Cambridge, Massachusetts

NASA Contract NAS 12-87

HIGH ORDER TIME DERIVATIVES OF POWERS OF THE RADIUS VECTOR *

By P. Sconzo and D. Valenzuela
IBM, Center for Exploratory Studies
Cambridge, Massachusetts

SUMMARY

The derivatives of a function of a function have been automatically computed applying the recursive Schlömilch-Cesàro formulation. Tables given for this purpose in collections of mathematical formulas were thus extended up to the eighteenth order.

Then, the general expressions thus obtained were used to construct the power series expansions in the time variable of those powers of the radius vector which appear most frequently in celestial mechanics, namely, the force function, the force of attraction, and the derivatives of the force function with respect to the coordinates. These last expressions are given in terms of the radius vector evaluated at origin and the triplet of Stumpff's local invariants. The explicit expressions of these series were symbolically computed using the FORMAC language.

INTRODUCTION

The need to compute the Taylor series expansion of a function of a function arises frequently in celestial mechanics. Even though by the standards of this discipline the computation of such series is an elementary problem, there is no doubt that its actual computation, if carried out by hand and to a high order, can be very tedious and time-consuming. This seems to be a case, therefore, where a computer programmed to perform symbol manipulation could relieve the scientists of this heavy task. We present here, first, a general algorithm given by Schlömilch and Cesàro to compute the derivatives of a function of a function in a recursive mode.

---

* This report presents part of the work being performed under contract NAS 12-87, "Automatic Symbol Processing Techniques."

This algorithm was programmed in FORMAC[1] and was used to compute the derivatives up to the order $n = 18$. Then, these general results were applied to obtain the expansions of $r^{-k}(t)$, where $k=1, 2, 3$.

FAÀ di BRUNO'S FORMULA

Let

$$u(t) = u[\, v(t)\,] \tag{1}$$

be the function which we want to expand in Taylor series:

$$u = \sum_{n=0}^{\infty} \frac{1}{n\,!} \left[\frac{d^n u}{d^t n} \right]_{t=0} t^n = \sum_{n=0}^{\infty} c_n t^n \tag{2}$$

Then, the goal of the computation is to obtain explicit expressions for the coefficients c_n. For small n they are

$$c_1 = \left[\frac{du}{dv} \cdot \frac{dv}{dt} \right]_{t=0}$$

$$c_2 = \frac{1}{2\,!} \left[\frac{d^2u}{dv^2} \cdot \left(\frac{dv}{dt} \right)^2 + \frac{du}{dv} \frac{d^2v}{dt^2} \right]_{t=0} \tag{3}$$

$$c_3 = \frac{1}{3\,!} \left[\frac{d^3u}{dv^3} \left(\frac{dv}{dt} \right)^3 + 3 \frac{d^2u}{dv^2} \frac{dv}{dt} \frac{d^2v}{dt^2} + \frac{du}{dv} \frac{d^3v}{dt^3} \right]_{t=0}$$

and it is immediately evident that the amount of work required to obtain higher order derivatives increases rapidly with the order n.

In spite of the complexity of this process, Faà di Bruno[2] has given an elegant and concise formula for the n^{th} derivative of a function of a function:

$$(4)\qquad \frac{d^n u}{dt^n} = \sum \frac{n!}{i!\, j! \ldots k!} \frac{d^p u}{dv^p} \left(\frac{1}{1!}\frac{dv}{dt}\right)^i \left(\frac{1}{2!}\frac{d^2 v}{dt^2}\right)^j \cdots \left(\frac{1}{l!}\frac{d^l v}{dt^l}\right)^k$$

where the summation sign extends over all the integer solutions of

$$(5)\qquad i + 2j + \ldots + lk = n$$

while, simultaneously, it is:

$$(6)\qquad i + j + \ldots + k = p.$$

The compendiousness of (4) makes it very adequate for theoretical work, for instance, to establish bounds, but it has little practical value for the actual computation of explicit expressions, particularly when high order derivatives are required. In effect, the most extended tabulations that can be used for this purpose, namely, those of the π- and M_3-numbers in table (24.2) of a well-known collection of mathematical formulas [3], give these numbers up to n = 10, while in two different applications to specific problems of celestial mechanics by Musen [4] in perturbation theory and Szebehely [5] in the restricted three-body problem, it has been found necessary to compute these explicit expressions up to n = 8. It seems likely that in future investigations a much higher order of approximation might be required.

SCHLÖMILCH-CESÀRO'S FORMULA

For $n > 10$, then, there is no way of avoiding a lengthy and tedious computation. This appears to be a case in which a computer, adequately programmed, could take the burden of the long literal developments thus relieving the mathematician from this uninteresting task.

While reviewing old and new mathematical literature on the subject of high-order derivatives of a function of a function, we have found a recursive procedure which adapts itself very well to being programmed in

FORMAC. We will call it the Schlömilch-Cesàro formulation because it can be found in the mathematical works of both authors[6]. According to this method,

$$\frac{d^n u}{dt^n} = \sum_{\nu=1}^{n} D_\nu \, P_{n,\nu} \tag{7}$$

where

$$D_\nu = \frac{d^\nu u}{dv^\nu} \tag{8}$$

and $P_{n,\nu}$ is a polynomial in the powers of

$$\alpha_i = \frac{d^i v}{dt^i} \tag{9}$$

Now, it is easy to establish the following recursion formula:

$$P_{n+1,\nu} = \frac{d}{dt} P_{n,\nu} + \alpha_1 P_{n,\nu-1} \qquad (\nu = 1, 2, \ldots, n+1) \tag{10}$$

which should be supplemented by

$$P_{n+1,0} = 0, \qquad P_{n+1,\nu} = 0 \qquad (\nu > n+1). \tag{11}$$

Noticing that for n = 1 there is only one polynomial

$$P_{1,1} = \alpha_1 = \frac{dv}{dt} , \tag{12}$$

we obtain successively for the first five orders:

$$
\begin{aligned}
P_{2,1} &= \alpha_2 \\
P_{2,2} &= \alpha_1^2
\end{aligned}
\qquad (n = 2)
$$

$$
\begin{aligned}
P_{3,1} &= \alpha_3 \\
P_{3,2} &= 3\alpha_1\ \alpha_2 \\
P_{3,3} &= \alpha_1^3
\end{aligned}
\qquad (n = 3)
$$

(13)
$$
\begin{aligned}
P_{4,1} &= \alpha_4 \\
P_{4,2} &= 4\alpha_1\alpha_3 + 3\alpha_2^2 \\
P_{4,3} &= 6\alpha_1^2\alpha_2 \\
P_{4,4} &= \alpha_1^4
\end{aligned}
\qquad (n = 4)
$$

$$
\begin{aligned}
P_{5,1} &= \alpha_5 \\
P_{5,2} &= 5\alpha_1\alpha_4 + 10\alpha_2\alpha_3 \\
P_{5,3} &= 10\alpha_1^2\alpha_3 + 15\alpha_1\alpha_2^2 \\
P_{5,4} &= 10\alpha_1^3\alpha_2 \\
P_{5,5} &= \alpha_1^5
\end{aligned}
\qquad (n = 5)
$$

This procedure has been used to generate $P_{n,\nu}$ up to $n = 18$, at which value the numeric coefficients become too large to be computed in exact integer form. The polynomials for the first ten orders were checked against the above mentioned tables [3] and were found to be identical.* As examples of the extension of these tables, we present

---

* Jordan [7] gives the number of terms in all the polynomials $P_{n,\nu}$ using Netto's notation $\Gamma(\int n)$. In the table on page 155 of reference [7], $\Gamma(\int 10)$ is given as 43, while our results and those of reference [3] indicate that the number of terms is 42.

in the appendix the polynomials $P_{11,\nu}$ and $P_{12,\nu}$.*

It is worth mentioning that these polynomials have a remarkable property: the sum of the coefficients is equal to the Stirling number of the second kind $S_n^{(\nu)}$. Thus, in the process of generating these polynomials, one also generates, as a by-product, the Stirling numbers.

APPLICATIONS

The method expressed in formulas (7) through (13) can be advantageously applied to computations frequently occurring in celestial mechanics. We show here with some detail its application to the construction of the time power series for the inverse square of the radius vector in the Keplerian motion along any conic section. In this case, r is a function of t through the Cartesian coordinates x, y, z. We take

$$(14) \qquad u = \frac{1}{r^2(t)}$$

Then, we have

$$(15) \qquad D_\nu = \frac{d^\nu u}{d r^\nu} = (-1)^\nu \frac{(\nu+1)!}{r^{\nu+2}}$$

Introducing Stumpff [8] local invariants μ, σ, ϵ we have the following expressions[9] for the time derivatives of r(t):

$$(16) \qquad \begin{aligned} \alpha_1 &= r_0 \sigma \\ \alpha_2 &= r_0(-\sigma^2 + \epsilon) \end{aligned}$$

---

* Polynomials of higher order ($n \leq 18$) can be requested from the authors.

$$
\begin{aligned}
\alpha_3 &= r_0(3\sigma^3 - \sigma\mu - 3\sigma\epsilon) \\
\alpha_4 &= r_0(-15\sigma^4 + 7\sigma^2\mu + 18\sigma^2\epsilon - \mu\epsilon - 3\epsilon^2) \qquad (16) \\
\alpha_5 &= r_0(105\sigma^5 - 60\sigma^3\mu - 150\sigma^3\epsilon + \sigma\mu^2 + 45\sigma\epsilon^2 + 24\mu\sigma\epsilon)
\end{aligned}
$$

where r_0 is the initial value of r.

The polynomials $P_{n,\nu}$ can, thus, be expressed in terms of the triplet μ, σ, ϵ substituting (16) into (13). Then, taking into account (15) it is easy to obtain

$$
\begin{aligned}
\frac{1}{1!}\left(\frac{du}{dt}\right)_0 &= -\frac{2}{r_0^2}\sigma \\
\frac{1}{2!}\left(\frac{d^2u}{dt^2}\right) &= \frac{1}{r_0^2}(4\sigma^2 - \epsilon) \\
\frac{1}{3!}\left(\frac{d^3u}{dt^3}\right) &= -\frac{1}{r_0^2}\left(8\sigma^3 - \frac{1}{3}\sigma\mu - 4\sigma\epsilon\right) \qquad (17) \\
\frac{1}{4!}\left(\frac{d^4u}{dt^4}\right) &= \frac{1}{r_0^2}\left(16\sigma^4 - \frac{19}{12}\sigma^2\mu - 12\sigma^2\epsilon + \frac{1}{12}\mu\epsilon + \epsilon^2\right) \\
\frac{1}{5!}\left(\frac{d^5u}{dt^5}\right) &= -\frac{1}{r_0^2}\left(32\sigma^5 - \frac{21}{4}\sigma^3\mu - 32\sigma^3\epsilon + \frac{1}{60}\sigma\mu^2 + 6\sigma\epsilon^2 + \frac{23}{20}\mu\sigma\epsilon\right).
\end{aligned}
$$

Expressions for terms of higher order are presented in the appendix, together with the corresponding terms for the inverse and inverse cube of the radius vector. All these expressions were produced by a single FORMAC program. The numeric coefficients were computed in exact form using rational arithmetic. These computations were truncated at n = 14, because overflow occurred while generating the numeric coefficients for the following term. These same computations have also been performed in floating point form; in this case it is possible to compute the first twenty terms of these series. The expansions of any power of the radius vector could be obtained using this program.

The time power series for r^{-2} can be used to obtain directly the true anomaly through the integral of areas for any value of the time within the radius of convergence. The series for r^{-1} and r^{-3} can be applied to obtain the power series solution of the three-body problem in the time domain.

The series for r^{-2} has also been applied[10] to the computation of ephemerides in the case of nearly parabolic orbits.

REFERENCES

1. E.R. Bond et al. FORMAC, an Experimental Formula Manipulation Compiler. Proceedings of the ACM National Conference, August 1964; IBM Technical Report 00.192, September 1964.

2. F. Faà di Bruno. Annali di Tortolini, Vol. 6, (1855). See also C.J. de la Vallée-Poussin, Course D Analyse Infinitesimale, Libr. Univ. Louvain, Vol. 1, p. 89, (1926).

3. Handbook of Mathematical Functions. National Bureau of Standards, Washington, D.C., (1964).

4. P. Musen. Journal of the Astronautical Sciences, Vol. 2, p. 129, (1965).

5. V. Szebehely. Astronomical Journal, Vol. 71, p. 968 (1966).

6. E. Cesàro. Corso di An. Algebrica. Ediz. Bocca Torino, p. 496, (1896).

7. C. Jordan. Calculus of Finite Differences. Chelsea Publish., New York (1950).

8. K. Stumpff. Himmelsmechanik. Deutscher Verlag, Vol. 1, p. 292, (1959).

9. P. Sconzo. Memorie della Società Astronomica Italiana, Vol. XXXVIII, p. 85-105, (1967). On page 104, the signs of the coefficients of $\sigma^2\mu^4$, $\sigma^6\epsilon\mu$, $\sigma^6\epsilon^2$ should be corrected to positive.

10. P. Sconzo. To be published in Zeitschrift angew. Math.

APPENDIX

Table 1

Polynomials $P_{11,\nu}$

$$P_{11,1} = \alpha_{11}$$

$$P_{11,2} = 11\alpha_1\alpha_{10} + 55\alpha_2\alpha_9 + 165\alpha_3\alpha_8 + 330\alpha_4\alpha_7 + 462\alpha_5\alpha_6$$

$$P_{11,3} = 55\alpha_1^2\alpha_9 + 495\alpha_1\alpha_2\alpha_8 + 1320\alpha_1\alpha_3\alpha_7 + 2310\alpha_1\alpha_4\alpha_6$$
$$+ 1386\alpha_1\alpha_5^2 + 990\alpha_2^2\alpha_7 + 4620\alpha_2\alpha_3\alpha_6 + 6930\alpha_2\alpha_4\alpha_5$$
$$+ 5775\alpha_3\alpha_4^2 + 4620\alpha_3^2\alpha_5$$

$$P_{11,4} = 165\alpha_1^3\alpha_8 + 1980\alpha_1^2\alpha_2\alpha_7 + 4620\alpha_1^2\alpha_3\alpha_6 + 6930\alpha_1^2\alpha_4\alpha_5$$
$$+ 17325\alpha_1\alpha_2\alpha_4^2 + 6930\alpha_1\alpha_2^2\alpha_6 + 27720\alpha_1\alpha_2\alpha_3\alpha_5$$
$$+ 23100\alpha_1\alpha_3^2\alpha_4 + 6930\alpha_2^3\alpha_5 + 34650\alpha_2^2\alpha_3\alpha_4 + 15400\alpha_2\alpha_3^3$$

$$P_{11,5} = 330\alpha_1^4\alpha_7 + 4620\alpha_1^3\alpha_2\alpha_6 + 9240\alpha_1^3\alpha_3\alpha_5 + 5775\alpha_1^3\alpha_4^2$$
$$+ 69300\alpha_1^2\alpha_2\alpha_3\alpha_4 + 20790\alpha_1^2\alpha_2^2\alpha_5 + 15400\alpha_1^2\alpha_3^3$$
$$+ 69300\alpha_1\alpha_2^2\alpha_3^2 + 34650\alpha_1\alpha_2^3\alpha_4 + 17325\alpha_2^4\alpha_3$$

$$P_{11,6} = 462\alpha_1^5\alpha_6 + 6930\alpha_1^4\alpha_2\alpha_5 + 11550\alpha_1^4\alpha_3\alpha_4 + 46200\alpha_1^3\alpha_2\alpha_3^2$$
$$+ 34650\alpha_1^3\alpha_2^2\alpha_4 + 69300\alpha_1^2\alpha_2^3\alpha_3 + 10395\alpha_1\alpha_2^5$$

$$P_{11,7} = 462\,\alpha_1^6\alpha_5 + 6930\alpha_1^5\alpha_2\alpha_4 + 4620\,\alpha_1^5\alpha_3^2 + 34650\,\alpha_1^4\alpha_2^2\alpha_3 + 17325\,\alpha_1^3\alpha_2^4$$

$$P_{11,8} = 330\,\alpha_1^7\alpha_4 + 4620\,\alpha_1^6\alpha_2\alpha_3 + 6930\,\alpha_1^5\alpha_2^3$$

$$P_{11,9} = 165\,\alpha_1^8\alpha_3 + 990\,\alpha_1^7\alpha_2^2$$

$$P_{11,10} = 55\,\alpha_1^9\alpha_2$$

$$P_{11,11} = \alpha_1^{11}$$

Table 2

Polynomials $P_{12,\nu}$

$$P_{12,1} = \alpha_{12}$$

$$P_{12,2} = 12\,\alpha_1\alpha_{11} + 66\,\alpha_2\alpha_{10} + 220\,\alpha_3\alpha_9 + 495\,\alpha_4\alpha_8 + 792\,\alpha_5\alpha_7 + 462\,\alpha_6^2$$

$$P_{12,3} = 66\,\alpha_1^2\alpha_{10} + 660\,\alpha_1\alpha_2\alpha_9 + 1980\,\alpha_1\alpha_3\alpha_8 + 3960\,\alpha_1\alpha_4\alpha_7 + 5544\,\alpha_1\alpha_5\alpha_6 + 1485\,\alpha_2^2\alpha_8 + 7920\,\alpha_2\alpha_3\alpha_7 + 8316\,\alpha_2\alpha_5^2 + 13860\,\alpha_2\alpha_4\alpha_6 + 9240\,\alpha_3^2\alpha_6 + 27720\,\alpha_3\alpha_4\alpha_5 + 5775\,\alpha_4^3$$

$$\begin{aligned}
P_{12,4} \; = \; & 220\,\alpha_1^3\alpha_9 + 2970\,\alpha_1^2\alpha_2\alpha_8 + 7920\,\alpha_1^2\alpha_3\alpha_7 + 13860\,\alpha_1^2\alpha_4\alpha_6 \\
& + 8316\,\alpha_1^2\alpha_5^2 + 55440\,\alpha_1\alpha_2\alpha_3\alpha_6 + 83160\,\alpha_1\alpha_2\alpha_4\alpha_5 \\
& + 11880\,\alpha_1\alpha_2^2\alpha_7 + 55440\,\alpha_1\alpha_3^2\alpha_5 + 69300\,\alpha_1\alpha_3\alpha_4^2 \\
& + 13860\,\alpha_2^3\alpha_6 + 83160\,\alpha_2^2\alpha_3\alpha_5 + 51975\,\alpha_2^2\alpha_4^2 \\
& + 138600\,\alpha_2\alpha_3^2\alpha_4 + 15400\,\alpha_3^4
\end{aligned}$$

$$\begin{aligned}
P_{12,5} \; = \; & 495\,\alpha_1^4\alpha_8 + 7920\,\alpha_1^3\alpha_2\alpha_7 + 18480\,\alpha_1^3\alpha_3\alpha_6 + 27720\,\alpha_1^3\alpha_4\alpha_5 \\
& + 166320\,\alpha_1^2\alpha_2\alpha_3\alpha_5 + 41580\,\alpha_1^2\alpha_2^2\alpha_6 + 103950\,\alpha_1^2\alpha_2\alpha_4^2 \\
& + 138600\,\alpha_1^2\alpha_3^2\alpha_4 + 184800\,\alpha_1\alpha_2\alpha_3^3 + 415800\,\alpha_1\alpha_2^2\alpha_3\alpha_4 \\
& + 83160\,\alpha_1\alpha_2^3\alpha_5 + 51975\,\alpha_2^4\alpha_4 + 138600\,\alpha_2^3\alpha_3^2
\end{aligned}$$

$$\begin{aligned}
P_{12,6} \; = \; & 792\,\alpha_1^5\alpha_7 + 13860\,\alpha_1^4\alpha_2\alpha_6 + 27720\,\alpha_1^4\alpha_3\alpha_5 + 17325\,\alpha_1^4\alpha_4^2 \\
& + 83160\,\alpha_1^3\alpha_2^2\alpha_5 + 277200\,\alpha_1^3\alpha_2\alpha_3\alpha_4 + 61600\,\alpha_1^3\alpha_3^3 \\
& + 415800\,\alpha_1^2\alpha_2^2\alpha_3^2 + 207900\,\alpha_1^2\alpha_2^3\alpha_4 + 207900\,\alpha_1\alpha_2^4\alpha_3 \\
& + 10395\,\alpha_2^6
\end{aligned}$$

$$\begin{aligned}
P_{12,7} \; = \; & 924\,\alpha_1^6\alpha_6 + 16632\,\alpha_1^5\alpha_2\alpha_5 + 27720\,\alpha_1^5\alpha_3\alpha_4 + 138600\,\alpha_1^4\alpha_2\alpha_3^2 \\
& + 103950\,\alpha_1^4\alpha_2^2\alpha_4 + 277200\,\alpha_1^3\alpha_2^3\alpha_3 + 62370\,\alpha_1^2\alpha_2^5
\end{aligned}$$

$$P_{12,8} = 792\,\alpha_1^7\alpha_5 + 13860\,\alpha_1^6\alpha_2\alpha_4 + 9240\,\alpha_1^6\alpha_3^2 + 83160\,\alpha_1^5\alpha_2^2\alpha_3 + 51975\,\alpha_1^4\alpha_2^4$$

$$P_{12,9} = 495\,\alpha_1^8\alpha_4 + 7920\,\alpha_1^7\alpha_2\alpha_3 + 13860\,\alpha_1^6\alpha_2^3$$

$$P_{12,10} = 220\,\alpha_1^9\alpha_3 + 1485\,\alpha_1^8\alpha_2^2$$

$$P_{12,11} = 66\,\alpha_1^{10}\alpha_2$$

$$P_{12,12} = \alpha_1^{12}$$

Table 3

n	$c_n = \frac{1}{n!}\,\frac{d^n}{dt^n}\left(\frac{1}{r}\right)$
1	$-\frac{1}{r}\sigma$
2	$\frac{1}{r}\left(-\frac{\epsilon}{2} + \frac{3}{2}\sigma^2\right)$
3	$\frac{1}{r}\left(\frac{3}{2}\epsilon\sigma + \frac{\mu}{6}\sigma - \frac{5}{2}\sigma^3\right)$
4	$\frac{1}{r}\left(\frac{\epsilon}{24}\mu - \frac{15}{4}\epsilon\sigma^2 + \frac{3}{8}\epsilon^2 - \frac{5}{8}\mu\sigma^2 + \frac{35}{8}\sigma^4\right)$
5	$\frac{1}{r}\left(-\frac{9}{20}\epsilon\mu\sigma + \frac{35}{4}\epsilon\sigma^3 - \frac{15}{8}\epsilon^2\sigma + \frac{7}{4}\mu\sigma^3 - \frac{\mu^2}{120}\sigma - \frac{63}{8}\sigma^5\right)$

6 $$\frac{1}{r}\left(\frac{21}{10}\epsilon\mu\sigma^2 - \frac{\epsilon}{720}\mu^2 - \frac{315}{16}\epsilon\sigma^4 - \frac{3}{40}\epsilon^2\mu + \frac{105}{16}\epsilon^2\sigma^2 - \frac{5}{16}\epsilon^3 - \frac{35}{8}\mu\sigma^4 + \frac{7}{80}\mu^2\sigma^2 + \frac{231}{16}\sigma^6\right)$$

7 $$\frac{1}{r}\left(-\frac{59}{8}\epsilon\mu\sigma^3 + \frac{27}{560}\epsilon\mu^2\sigma + \frac{693}{16}\epsilon\sigma^5 + \frac{459}{560}\epsilon^2\mu\sigma - \frac{315}{16}\epsilon^2\sigma^3 + \frac{35}{16}\epsilon^3\sigma + \frac{165}{16}\mu\sigma^5 - \frac{7}{16}\mu^2\sigma^3 + \frac{\mu^3}{5040}\sigma - \frac{429}{16}\sigma^7\right)$$

8 $$\frac{1}{r}\left(\frac{2871}{128}\epsilon\mu\sigma^4 - \frac{195}{448}\epsilon\mu^2\sigma^2 + \frac{\epsilon}{40320}\mu^3 - \frac{3003}{32}\epsilon\sigma^6 - \frac{4131}{896}\epsilon^2\mu\sigma^2 + \frac{27}{4480}\epsilon^2\mu^2 + \frac{3465}{64}\epsilon^2\sigma^4 + \frac{459}{4480}\epsilon^3\mu - \frac{315}{32}\epsilon^3\sigma^2 + \frac{35}{128}\epsilon^4 - \frac{3003}{128}\mu\sigma^6 + \frac{209}{128}\mu^2\sigma^4 - \frac{17}{2688}\mu^3\sigma^2 + \frac{6435}{128}\sigma^8\right)$$

9 $$\frac{1}{r}\left(-\frac{1001}{16}\epsilon\mu\sigma^5 + \frac{1199}{504}\epsilon\mu^2\sigma^3 - \frac{\epsilon}{360}\mu^3\sigma + \frac{6435}{32}\epsilon\sigma^7 + \frac{4345}{224}\epsilon^2\mu\sigma^3 - \frac{929}{6720}\epsilon^2\mu^2\sigma - \frac{9009}{64}\epsilon^2\sigma^5 - \frac{635}{504}\epsilon^3\mu\sigma + \frac{1155}{32}\epsilon^3\sigma^3 - \frac{315}{128}\epsilon^4\sigma + \frac{5005}{96}\mu\sigma^7 - \frac{1001}{192}\mu^2\sigma^5 + \frac{11}{189}\mu^3\sigma^3 - \frac{\mu^4}{362880}\sigma - \frac{12155}{128}\sigma^9\right)$$

10 $$\frac{1}{r}\left(\frac{79079}{480}\epsilon\mu\sigma^6 - \frac{1287}{128}\epsilon\mu^2\sigma^4 + \frac{4939}{100800}\epsilon\mu^3\sigma^2 - \frac{\epsilon}{3628800}\mu^4 - \frac{109395}{256}\epsilon\sigma^8 - \frac{4433}{64}\epsilon^2\mu\sigma^4 + \frac{28479}{22400}\epsilon^2\mu^2\sigma^2\right.$$

$$- \frac{\epsilon^2}{3600} \mu^3 + \frac{45045}{128} \epsilon^2 \sigma^6 + \frac{1397}{168} \epsilon^3 \mu \sigma^2 - \frac{929}{67200} \epsilon^3 \mu^2$$

$$- \frac{15015}{128} \epsilon^3 \sigma^4 - \frac{127}{1008} \epsilon^4 \mu + \frac{3465}{256} \epsilon^4 \sigma^2 - \frac{63}{256} \epsilon^5$$

$$- \frac{7293}{64} \mu \sigma^8 + \frac{29029}{1920} \mu^2 \sigma^6 - \frac{143}{432} \mu^3 \sigma^4 + \frac{341}{1209600} \mu^4 \sigma^2$$

$$+ \frac{46189}{256} \sigma^{10} \Big)$$

Table 4

n	$c_n = \frac{1}{n!} \frac{d^n}{dt^n} \left(\frac{1}{r^2} \right)$
1	$- \frac{2}{r^2} \sigma$
2	$\frac{1}{r^2} \left(- \epsilon + 4 \sigma^2 \right)$
3	$\frac{1}{r^2} \left(4 \epsilon \sigma + \frac{\mu}{3} \sigma - 8 \sigma^3 \right)$
4	$\frac{1}{r^2} \left(\frac{\epsilon}{12} \mu - 12 \epsilon \sigma^2 + \epsilon^2 - \frac{19}{12} \mu \sigma^2 + 16 \sigma^4 \right)$
5	$\frac{1}{r^2} \left(- \frac{23}{20} \epsilon \mu \sigma + 32 \epsilon \sigma^3 - 6 \epsilon^2 \sigma + \frac{21}{4} \mu \sigma^3 - \frac{\mu^2}{60} \sigma - 32 \sigma^5 \right)$
6	$\frac{1}{r^2} \Big(\frac{127}{20} \epsilon \mu \sigma^2 - \frac{\epsilon}{360} \mu^2 - 80 \epsilon \sigma^4 - \frac{23}{120} \epsilon^2 \mu + 24 \epsilon^2 \sigma^2$ $- \epsilon^3 - \frac{359}{24} \mu \sigma^4 + \frac{79}{360} \mu^2 \sigma^2 + 64 \sigma^6 \Big)$

7 $$\frac{1}{r^2}\Big(-\frac{1523}{60}\epsilon\mu\sigma^3+\frac{17}{140}\epsilon\mu^2\sigma+192\epsilon\sigma^5+\frac{697}{280}\epsilon^2\mu\sigma$$
$$-80\epsilon^2\sigma^3+8\epsilon^3\sigma+\frac{941}{24}\mu\sigma^5-\frac{77}{60}\mu^2\sigma^3+\frac{\mu^3}{2520}\sigma$$
$$-128\sigma^7\Big)$$

8 $$\frac{1}{r^2}\Big(\frac{27441}{320}\epsilon\mu\sigma^4-\frac{9}{7}\epsilon\mu^2\sigma^2+\frac{\epsilon}{20160}\mu^3-448\epsilon\sigma^6$$
$$-\frac{35709}{2240}\epsilon^2\mu\sigma^2+\frac{17}{1120}\epsilon^2\mu^2+240\epsilon^2\sigma^4+\frac{697}{2240}\epsilon^3\mu$$
$$-40\epsilon^3\sigma^2+\epsilon^4-\frac{6241}{64}\mu\sigma^6+\frac{867}{160}\mu^2\sigma^4-\frac{319}{20160}\mu^3\sigma^2$$
$$+256\sigma^8\Big)$$

9 $$\frac{1}{r^2}\Big(-\frac{16721}{64}\epsilon\mu\sigma^5+\frac{1781}{224}\epsilon\mu^2\sigma^3-\frac{421}{60480}\epsilon\mu^3\sigma$$
$$+1024\epsilon\sigma^7+\frac{100091}{1344}\epsilon^2\mu\sigma^3-\frac{2753}{6720}\epsilon^2\mu^2\sigma-672\epsilon^2\sigma^5$$
$$-\frac{17609}{4032}\epsilon^3\mu\sigma+160\epsilon^3\sigma^3-10\epsilon^4\sigma+\frac{44923}{192}\mu\sigma^7$$
$$-\frac{3679}{192}\mu^2\sigma^5+\frac{2047}{12096}\mu^3\sigma^3-\frac{\mu^4\sigma}{181440}-512\sigma^9\Big)$$

10 $$\frac{1}{r^2}\Big(\frac{2227}{3}\epsilon\mu\sigma^6-\frac{166641}{4480}\epsilon\mu^2\sigma^4+\frac{14429}{100800}\epsilon\mu^3\sigma^2$$
$$-\frac{\epsilon\mu^4}{1814400}-2304\epsilon\sigma^8-\frac{130073}{448}\epsilon^2\mu\sigma^4$$
$$+\frac{286877}{67200}\epsilon^2\mu^2\sigma^2-\frac{421}{604800}\epsilon^2\mu^3+1792\epsilon^2\sigma^6$$
$$+\frac{161377}{5040}\epsilon^3\mu\sigma^2-\frac{2753}{67200}\epsilon^3\mu^2-560\epsilon^3\sigma^4-\frac{17609}{40320}\epsilon^4\mu$$
$$+60\epsilon^4\sigma^2-\epsilon^5-\frac{210029}{384}\mu\sigma^8+\frac{7759}{128}\mu^2\sigma^6$$

$$- \frac{130973}{120960} \mu^3 \sigma^4 + \frac{1279}{1814400} \mu^4 \sigma^2 + 1024 \sigma^{10} \Big)$$

Table 5

n	$c_n = \frac{1}{n!} \frac{d^n}{dt^n} \left(\frac{1}{r^3} \right)$
1	$- \frac{3}{r^3} \sigma$
2	$\frac{1}{r^3} \left(- \frac{3}{2} \epsilon + \frac{15}{2} \sigma^2 \right)$
3	$\frac{1}{r^3} \left(\frac{15}{2} \epsilon \sigma + \frac{\mu}{2} \sigma - \frac{35}{2} \sigma^3 \right)$
4	$\frac{1}{r^3} \left(\frac{\epsilon}{8} \mu - \frac{105}{4} \epsilon \sigma^2 + \frac{15}{8} \epsilon^2 - \frac{23}{8} \mu \sigma^2 + \frac{315}{8} \sigma^4 \right)$
5	$\frac{1}{r^3} \left(- \frac{21}{10} \epsilon \mu \sigma + \frac{315}{4} \epsilon \sigma^3 - \frac{105}{8} \epsilon^2 \sigma + 11 \mu \sigma^3 - \frac{\mu^2}{40} \sigma - \frac{693}{8} \sigma^5 \right)$
6	$\frac{1}{r^3} \left(\frac{107}{8} \epsilon \mu \sigma^2 - \frac{\epsilon}{240} \mu^2 - \frac{3465}{16} \epsilon \sigma^4 - \frac{7}{20} \epsilon^2 \mu + \frac{945}{16} \epsilon^2 \sigma^2 - \frac{35}{16} \epsilon^3 - \frac{281}{8} \mu \sigma^4 + \frac{19}{48} \mu^2 \sigma^2 + \frac{3003}{16} \sigma^6 \right)$
7	$\frac{1}{r^3} \Big(- \frac{479}{8} \epsilon \mu \sigma^3 + \frac{123}{560} \epsilon \mu^2 \sigma + \frac{9009}{16} \epsilon \sigma^5 + \frac{589}{112} \epsilon^2 \mu \sigma - \frac{3465}{16} \epsilon^2 \sigma^3 + \frac{315}{16} \epsilon^3 \sigma + \frac{1619}{16} \mu \sigma^5 - \frac{127}{48} \mu^2 \sigma^3$

$$+ \frac{\mu^3}{1680}\sigma - \frac{6435}{16}\sigma^7 \Big)$$

8 $$\frac{1}{r^3}\Big(\frac{28437}{128}\epsilon\mu\sigma^4 - \frac{5967}{2240}\epsilon\mu^2\sigma^2 + \frac{\epsilon}{13440}\mu^3$$

$$- \frac{45045}{32}\epsilon\sigma^6 - \frac{33801}{896}\epsilon^2\mu\sigma^2 + \frac{123}{4480}\epsilon^2\mu^2$$

$$+ \frac{45045}{64}\epsilon^2\sigma^4 + \frac{589}{896}\epsilon^3\mu - \frac{3465}{32}\epsilon^3\sigma^2 + \frac{315}{128}\epsilon^4$$

$$- \frac{34913}{128}\mu\sigma^6 + \frac{1593}{128}\mu^2\sigma^4 - \frac{383}{13440}\mu^3\sigma^2 + \frac{109395}{128}\sigma^8 \Big)$$

9 $$\frac{1}{r^3}\Big(- \frac{46925}{64}\epsilon\mu\sigma^5 + \frac{881}{48}\epsilon\mu^2\sigma^3 - \frac{253}{20160}\epsilon\mu^3\sigma$$

$$+ \frac{109395}{32}\epsilon\sigma^7 + \frac{12385}{64}\epsilon^2\mu\sigma^3 - \frac{5717}{6720}\epsilon^2\mu^2\sigma$$

$$- \frac{135135}{64}\epsilon^2\sigma^5 - \frac{13915}{1344}\epsilon^3\mu\sigma + \frac{15015}{32}\epsilon^3\sigma^3 - \frac{3465}{128}\epsilon^4\sigma$$

$$+ \frac{44923}{64}\mu\sigma^7 - \frac{9253}{192}\mu^2\sigma^5 + \frac{599}{1728}\mu^3\sigma^3 - \frac{\mu^4}{120960}\sigma$$

$$- \frac{230945}{128}\sigma^9 \Big)$$

10 $$\frac{1}{r^3}\Big(\frac{1429381}{640}\epsilon\mu\sigma^6 - \frac{60161}{640}\epsilon\mu^2\sigma^4 + \frac{11863}{40320}\epsilon\mu^3\sigma^2$$

$$- \frac{\epsilon}{1209600}\mu^4 - \frac{2078505}{256}\epsilon\sigma^8 - \frac{104575}{128}\epsilon^2\mu\sigma^4$$

$$+ \frac{2215}{224}\epsilon^2\mu^2\sigma^2 - \frac{253}{201600}\epsilon^2\mu^3 + \frac{765765}{128}\epsilon^2\sigma^6$$

$$+ \frac{32017}{384}\epsilon^3\mu\sigma^2 - \frac{5717}{67200}\epsilon^3\mu^2 - \frac{225225}{128}\epsilon^3\sigma^4 - \frac{2783}{2688}\epsilon^4\mu$$

$$+ \frac{45045}{256}\epsilon^4\sigma^2 - \frac{693}{256}\epsilon^5 - \frac{111725}{64}\mu\sigma^8 + \frac{158291}{960}\mu^2\sigma^6$$

$$- \frac{787}{320}\mu^3\sigma^4 + \frac{307}{241920}\mu^4\sigma^2 + \frac{969969}{256}\sigma^{10} \Big)$$

ON THE LONG-PERIOD BEHAVIOUR OF CLOSE LUNAR ORBITERS

By J. Vagners
University of Washington
Seattle, Washington

NASA Research Grant NsG 133-61

* * * * * * *

ON THE LONG-PERIOD BEHAVIOUR OF CLOSE LUNAR ORBITERS*

By J. Vagners**

University of Washington
Seattle, Washington

SUMMARY

* * * * * * *

In this paper semi-analytical results are presented for the long-term behaviour of close lunar orbiters. The Moon's aspherical gravity field is described by spherical harmonics through $J_{4,4}$ and the Earth is idealized as moving in a circle in the lunar equatorial plane. After the short- and medium-period terms have been removed from the Hamiltonian, the long-period motion is analyzed from equi-energy trajectories in the eccentricity-argument of pericenter plane. Stationary, or resonant, solutions to the slowly varying equations of motion are determined numerically and the results presented as variations of (critical) inclination and eccentricity with semi-major axis. These computations are based on two sets of preliminary values of the lunar harmonics $J_{n,m}$ recently published by sources in the United States and the Soviet Union. Representative equi-energy contours are presented to illustrate the evolution of the long-term motion as influenced by orbital inclination, semi-major axis and the parameters $J_{n,m}$. Many stable orbits, in the sense of not impacting the Moon, are found even for high inclinations.

Introduction

In the analysis of near-earth satellite motion the dominant perturbation is the oblateness J_2, of $O(10^{-3})$‡, with all other (gravitational) perturbations of at most $O(10^{-6})$, or second order in J_2. As can be shown(1), to first order the Hamiltonian contains only short-period (periodic in the mean anomaly) and secular terms. These characteristics allow an analytical solution by successive approximation wherein, to any desired order of accuracy, the equations of motion are reduced to trivial quadratures.

This situation changes significantly when we consider the motion of a close lunar orbiter. The ratio of the mass of the Earth to the mass of the Moon is roughly 81 to 1. Consequently, when the orbits of a lunar orbiter and Earth orbiter are geometrically equivalent, the perturbation of the Earth will affect the lunar satellite 81^2 times stronger than the Moon will affect the orbit of the Earth satellite. From the early lunar orbiter flights it is known that the oblateness of the Moon is of $O(10^{-4})$ and that the higher harmonics $J_{n,m}$ in the lunar potential may be as large as $O(10^{-5})$. Hence, neither the Earth influences

---

* This work was performed in association with research sponsored by the National Aeronautics and Space Administration under Research Grant NsG 133-61.

**The research was performed while the author was at Stanford University and is a part of his Ph.D. Dissertation.

‡ The central force field is taken as $O(1)$.

nor the higher harmonics are of second order in J_2. In fact, as the semi-major axis of the satellite increases above two lunar radii, the Earth influences exceed, and eventually dominate, the oblateness effects. The result is that the variable part of the (slowly-varying) Hamiltonian is factored by small parameters of (roughly) the same order of magnitude so that successive approximation schemes fail. Nevertheless, several of the approximation steps may still be taken and then many significant features of the long-term motion extracted without actually integrating the equations of motion.

In modeling the lunar orbiter problem, I will make the following assumptions:

- The Earth moves in an apparent circular orbit in the lunar equatorial plane; the actual apparent orbit is inclined about 6°41' (±9') to the lunar equator and has an eccentricity of about 0.055.
- The Earth is spherically symmetrical and hence can be represented by a point mass.
- The lunar gravity field is described by spherical harmonics through $J_{4,4}$; at the time of this paper no estimates of the higher harmonics were available.
- Solar radiation pressure and solar gravity field effects can be neglected.
- The physical librations of the Moon can be ignored.

Since for close orbiters the neglected terms contribute small $O(10^{-7})$ perturbations or less, the following analysis should provide the correct gross behaviour. The only exceptions might prove to be the third and fourth assumptions; radiation pressure probably significantly affects the results if one considers orbiters of very high area-to-mass ratio, for example dust. The other, more general, invalidating factor might be the higher zonal harmonics J_n, $n \geq 5$.

Long Period Equations of Motion

In keeping with the assumptions on the apparent Earth orbit, I will take the reference plane to be the lunar equatorial plane with the zero-meridian toward the Earth. In terms of luni-centric coordinates the potential field of the Earth is expressed as follows

$$V_E = \frac{\mu_E}{r_E}\left[1 + \left(\frac{r}{r_E}\right)^2 P_2(s)\right] \tag{1}$$

where μ_E is the gravitational constant of the Earth, r_E is the (constant) luni-centric distance to the Earth, s is the direction cosine of the luni-centric angle between the satellite and the Earth, and P_2 is the second Legendre polynomial. The direction cosine s can be expressed in terms of orbital elements as

$$s = \cos(u + \Omega - \Omega_E) - 2\sin^2 i/2 \sin u \sin(\Omega_E - \Omega) \tag{2}$$

where Ω is the longitude of the ascending node, Ω_E is the mean longitude of the Earth, u is the central angle and i the inclination.

Following the device of Kozai[2], I define a set of modified Delaunay variables as follows:

$$L = \sqrt{\mu_{☾} a} \qquad \ell = \text{mean anomaly}$$
$$G = L\sqrt{1 - e^2} \qquad g = \text{argument of pericenter} \tag{3}$$
$$H = G \cos i \qquad h = \Omega - \Omega_E$$

where ℓ,g,h are canonically conjugate to L,G,H and e,a are the eccentricity and semi-major axis respectively. The associated Hamiltonian is

$$\mathcal{H} = - \frac{\mu_{☾}^2}{2L^2} - n_E H + \sum_{n,m} \mathcal{H}_{n,m} + V_E \tag{4}$$

where n_E is the (apparent) Earth mean motion. The term $\sum_{n,m} \mathcal{H}_{n,m}$ incorporates all of the lunar gravity anomalies described by the gravity coefficients $J_{n,m}$. The complete explicit form for all n,m is not of direct interest here; it may be found in Refs. 3 and 4. Under the assumption that the Earth orbit is circular and in the lunar equator, no angular variables other than ℓ,g,h appear in $\mathcal{H}$. Inclusion of Earth eccentricity and inclination would introduce angles such as the node and the argument of pericenter of the Earth orbit.

In general, the Hamiltonian (4) will contain the following types of terms: a) secular terms, depending only on L,G,H due to even zonal harmonics and the Earth, b) long-period terms (periodic in g) due to all zonal harmonics and the Earth, c) medium-period terms, with periods of a month or fractions thereof, arising from the tesseral harmonics and the Earth, and finally d) short- or ℓ-periodic terms from all sources. For the present problem the short- and medium-period terms may be removed from the Hamiltonian via the von Zeipel method[1]. The process will not be given here, since it is algebraically involved, although straight-forward, and may be found in Oesterwinter[5] and Giacaglia[6]. It may be noted here, however, that the referenced results are incomplete in view of the presently estimated magnitudes of the higher $J_{n,m}$ $[O(10^{-5})]$. In the quoted papers only the J_2, $J_{2,2}$ and Earth short-period and only $J_{2,2}$ and Earth medium-period terms were specifically determined. For the present purposes, I assume that all of the short- and medium-period terms have been removed leaving only a slowly varying Hamiltonian.

In the literature, the roles of the canonical coordinates and momenta are interchanged by considering the negative of the Hamiltonian, denoted by $\mathcal{F}$, and hence referring to $\mathcal{F}$ as the Hamiltonian. I shall adhere to this (by now) well-established convention. With the understanding that all elements are slowly varying, the Hamiltonian can be written as [4]:

$$\mathcal{F}^* = \frac{\mu_{☾}^2}{2L^2} + n_E H + \frac{1}{4} J_2 \frac{\mu_{☾}^4}{L^3 G^3}\left(1 - 3\frac{H^2}{G^2}\right) + \frac{n_E^2 L^4}{16\,\mu_{☾}^2}\left[\left(5 - 3\frac{G^2}{L^2}\right)\left(3\frac{H^2}{G^2} - 1\right)\right.$$
$$\left. + 15\left(1 - \frac{G^2}{L^2}\right)\left(1 - \frac{H^2}{G^2}\right)\right] - \frac{3 J_4 \mu_{☾}^6}{64 G^7 L^3}\left[(35 \sin^4 i - 40 \sin^2 i + 8)\left(1 + \frac{3e^2}{2}\right)\right.$$

$$-5\,(1 - 7\cos^2 i)\, e^2 \sin^2 i\Big] - \frac{3\, J_3 \mu_{☾}^5}{8 G^5 L^3}\; e \sin i\,(1 - 5\cos^2 i)\sin g$$

$$-\left[\frac{15\,\mu_{☾}^6 J_4}{32\, G^7 L^3}\; e^2 \sin^2 i\,(1 - 7\cos^2 i) + \frac{15\, n_E^2 L^4}{8\,\mu_{☾}^2}\left(1 - \frac{G^2}{L^2}\right)\left(1 - \frac{H^2}{G^2}\right)\right]\sin^2 g \tag{5}$$

In Eq. 5 the Kepler elements e and i have been used as a convenient shorthand for the more involved canonical counterparts and the mean lunar radius R = 1738 km has been taken as the unit of length.

Since $\mathcal{F}^*$ is independent of time, it is a constant as are L and H since neither ℓ nor h appears in $\mathcal{F}^*$. Thus we have a single degree of freedom system described by the differential equations

$$\dot{g} = -\frac{\partial \mathcal{F}^*}{\partial G}$$
$$\dot{G} = \frac{\partial \mathcal{F}^*}{\partial g} \tag{6}$$

Since the integral $\mathcal{F}^*$ = constant is also known, the problem, in principle at least, is solved. The known integral could be used to eliminate one of the variables and hence g,G would follow by quadrature. However, it can easily be ascertained that the solution cannot be found in terms of known functions. In the von Zeipel treatment(1) of a near-earth orbiter whose inclination is not to close to critical, the long-period fluctuations were removed from $\mathcal{F}^*$ at this point by a (further) canonical transformation and the problem reduced to trivial quadratures. Such a procedure fails in the present case as both the secular and g-dependent parts of the Hamiltonian have (essentially) an $O(10^{-4})$ multiplier. This can be deduced from (5) by noting that the multiplier of the Earth contribution, containing secular and cos 2g terms, has the factor‡ $(n_E/n')^2$ after division by $\mu_{☾}^2/2L^2$; this factor is a quantity of roughly the same order as J_2.

It proves convenient for numerical calculations to introduce the parameter η and constants of the motion β,α_i, defined by:

$$\eta = \frac{G}{L} = (1 - e^2)^{1/2} \quad , \quad \beta = \frac{H^2}{L^2} = (1 - e^2)\cos^2 i \tag{7}$$

and

$$\alpha_2 = \frac{J_2}{2a^2}\left(\frac{n'}{n_E}\right)^2$$

$$\alpha_3 = \frac{3J_3}{4a^3}\left(\frac{n'}{n_E}\right)^2$$

‡ $n^1 = \dfrac{\mu_{☾}^2}{L^3}$, the satellite mean motion.

$$\alpha_4 = \frac{3\ J_4}{16a^4}\left(\frac{n'}{n_E}\right)^2 \tag{8}$$

which are dimensionless since $R_{☾}$ has been chosen as the unit of length. Combining the constant parts (independent of g and G) with $\mathcal{F}^*$ and some algebraic manipulation yields the "energy integral" in the form:

$$\begin{aligned} C = {} & \alpha_2\eta^{-3}(1 - 3\beta\eta^{-2}) + \tfrac{1}{4}(5 + 3\beta - 6\eta^2) - \tfrac{1}{2}\,\alpha_4\eta^{-7}\left\{\tfrac{1}{2}\left[35(1 - \beta\eta^{-2})^2\right.\right. \\ & \left.\left. -40(1 - \beta\eta^{-2}) + 8\right](5 - 3\eta^2) - 5(1 - 7\beta\eta^{-2})(1 - \eta^2)(1 - \beta\eta^{-2})\right\} \\ & -\alpha_3\eta^{-5}(1 - \eta^2)^{1/2}\,(1 - \beta\eta^{-2})^{1/2}\,(1 - 5\beta\eta^{-2})\ \sin g - \frac{15}{4}\,(1 - \eta^2) \\ & \cdot\,(1 - \beta\eta^{-2})\left[\frac{4}{3}\,\alpha_4\,\eta^{-7}\,(1 - 7\beta\eta^{-2}) + 1\right]\ \sin^2 g \end{aligned} \tag{9}$$

where

$$C = \frac{2\mu_{☾}^2}{n_E^2 L^4}\left[\mathcal{F}^* - \frac{\mu_{☾}^2}{2L^2} - n_E H\right]$$

The form chosen here is such that if α_3 and α_4 are taken as zero, one recovers (essentially) the results of Kozai(2).

The equations of motion corresponding to this energy integral are determined by differentiation of (5) according to (6):

$$\begin{aligned} \dot{\eta} = \frac{1}{L}\,\dot{G} = \frac{1}{L}\,\frac{\partial\mathcal{F}^*}{\partial g} = -\frac{3\mu_{☾}}{8L}\left\{\frac{J_3}{a^4\eta^5}\left[(1 - \eta^2)(1 - \beta\eta^{-2})\right]^{1/2}(1 - 5\beta\eta^{-2})\ \cos g\right. \\ \left. -5(1 - \eta^2)(1 - \beta\eta^{-2})\left[\frac{J_4}{4a^5\eta^7}\ (1 - 7\beta\eta^{-2}) + \frac{n_E^2\, a^2}{\mu_{☾}}\right]\sin 2g\right\} \end{aligned} \tag{10}$$

$$\begin{aligned} \dot{g} = -\frac{\partial\mathcal{F}^*}{\partial G} = \frac{3}{4}\,\frac{\mu_{☾}}{L}\left\{\left\{\frac{J_2}{a^3\eta^6}\ (\eta^2 - 5\beta) + \frac{2n_E^2 a^2}{\mu_{☾}}\,\eta - \frac{5J_4}{32a^5\eta^{12}}\left[\eta^6 + 7(2\beta + 1)\eta^4\right.\right.\right. \\ \left.\left. - 63(\beta^2 + 2\beta)\eta^2 - 231\beta^2\right] + \frac{J_3}{2a^4\eta^5}\left\{-\frac{5}{\eta^4}\ (\eta^2 - 5\beta)\left[(1 - \eta^2)(1 - \beta\eta^{-2})\right]^{1/2}\right. \end{aligned}$$

$$- \frac{(\eta^2 - 5\beta)}{\eta^2} \left[\frac{\eta^2 - \beta}{1 - \eta^2}\right]^{1/2} + \beta \frac{(11\eta^2 - 15\beta)}{\eta^4} \left[\frac{1 - \eta^2}{\eta^2 - \beta}\right]^{1/2} \Bigg\} \sin g + \Bigg\{ \frac{5J_4}{8a^5\eta^{12}}$$

$$\cdot \left[5\eta^6 - (56\beta + 7)\eta^4 + (63\beta^2 + 72\beta)\eta^2 - 77\beta^2\right] + \frac{5n_E^2a^2}{\mu_{☾}\eta^3} (\beta - \eta^4) \Bigg\} \sin^2 g \Bigg\}\Bigg\} \quad (11)$$

Geometrical Solution

Since it seems highly unlikely that a general solution in terms of known functions can be found to the system (10) and (11), I will attempt to obtain the general characteristics of the motion indirectly. From the known energy integral (9) it is possible to construct, with the aid of a computer, equi-energy contours in the $\eta^2 - g$ plane for given values of the semi-major axis a and (angular momentum) parameter β. The form of the contours is also strongly dependent, for low a, on the values of the harmonic coefficients J_2, J_3 and J_4; as noted earlier these values are not very well known at the time of this writing.

From preliminary analysis of the U.S. lunar orbiter data‡ the following coefficients are available

$$J_2 = -2.3691 \times 10^{-4}$$

$$J_3 = 3.366 \times 10^{-5}$$

$$J_4 = -1.368 \times 10^{-5} \quad (12)$$

Another (preliminary) set, as determined from the U.S.S.R. Luna 10 flight[7] is

$$J_2 = -2.06 \times 10^{-4}$$

$$J_3 = -3.62 \times 10^{-5}$$

$$J_4 = 3.33 \times 10^{-5} \quad (13)$$

The agreement as to the value of J_2 is not too bad; the U.S. value is closer to that determined from the physical libration of the Moon (-2.41×10^{-4}) than the

‡Results presented at Guidance Theory and Trajectory Analysis Seminar, NASA Electronics Research Center, Cambridge, Mass., June 1967, by W. T. Blackshear, et al. of Langley Research Center.

U.S.S.R. value. As for J_3, the signs are different, although the magnitudes are fairly close, as is also the case with J_4. In view of these facts both sets of coefficients were used in determining typical "phase plane" (η^2, g) contours, thus in some sense indicating the sensitivity to, as well as dependence on, the values of the J_3 and J_4 harmonics.

With the motion constrained to a given energy curve, by plotting equi-energy contours we can obtain the qualitative and quantitative long-period behavior of g and eccentricity (or inclination). The time history on the contours can be obtained very rapidly via numerical integration of (10) and (11) for any contour of interest.

Before presenting some representative contours, note that for each value of the semi-major axis there is a maximum allowable eccentricity in order that the radius of pericenter remains greater than the radius of the Moon. This condition is simply $e_{max} = \frac{a-1}{a}$, or in terms of η^2

$$\eta^2_{crit} = \frac{2a - 1}{a^2} \tag{14}$$

For convenience, this relationship is shown graphically in Fig. 1. All equi-energy contours in the η^2 - g plane are terminated at η^2_{crit} since the crossing of the η^2_{crit} line by any contour implies lunar surface impact.

In Figs. 2 through 7 typical η^2 - g contours are presented; Figs. 2, 3, 4 were obtained using the Langley J_n values, Figs. 5 and 6 were obtained for the U.S.S.R. values. In Fig. 7 the set of elements corresponding to the U.S. Lunar Orbiter IV initial elements was chosen and the corresponding trajectories were also computed via numerical integration of the long-period equations of motion. The numerical integration points are given at 30-day intervals to indicate the relative rates of η^2, g on various parts of the contours. The maximum possible lifetime available for the parameters of Fig. 7 is roughly 3.5 years; the appropriate contour is indicated on the figure.

Inspection of the contours reveals several major features:

- Stationary points with surrounding librating orbits where g is constrained to vary between (some) limits.
- Circulating orbits where g increases through 2π.
- Impacting orbits resulting from either circulating or librating orbits.

From these observations and Figs. 2-7, one can deduce that stable, or surviving orbits are possible even for high inclinations if the initial conditions are chosen (sufficiently) near one of the stationary solutions.

I will, therefore, defer some additional comments concerning the η^2 - g contours until the nature of the stationary solutions has been investigated in the next section. Only stationary points for libration orbits are depicted in Figs. 2-7; one might ask if there are any other types of stationary solutions possible. Stationary, "saddle-type" points are also possible, but were found to exist only for η^2 less than η^2_{crit}, i.e., under the surface of the Moon. An example of contours illustrating the different types of stationary points and contours is given in Fig. 8. In this figure only the interval $0° \leq g \leq 90°$ is shown; the contours for $-90° \leq g \leq 0°$ are similar but not symmetrical about $g = 0°$.

Stationary Solutions

The stationary solutions are defined by $\dot{\eta} = \dot{g} = 0$ and correspond to resonances in the critical inclination sense, i.e., commensurabilities of the anomalistic and nodal periods. In the near-earth satellite problem where J_2 dominated all other perturbations, and hence the nodal regression rate, one found resonance at the critical inclination of ±63.4°. As the semi-major axis of the orbiter is increased and the luni-solar perturbation effects increase, coupled with the decreasing rate of the node due to J_2, one finds additional resonances emerging. For example, in an investigation of high-eccentricity, sun-perturbed orbits around Mars, Breakwell and Hensley(8) find eleven critical inclinations (resonances), all above $i = 40°$. The analysis of Breakwell and Hensley is inapplicable in the present case since it assumed that J_2 of Mars dominated the Sun perturbations, whereas in the lunar orbiter problem J_2, J_3, J_4 and $(n_E/n')^2$ are all roughly the same for close lunar orbiters. Characteristic of these resonances are the associated large fluctuations in eccentricity and the inclination.

Digressing, we note that the physical reality of such resonances was perhaps first verified by Explorer VI, launched August 1959, whose elements were $i = 47.1$, $a = 4.35\ R_E$, $e = 0.76$, placing it very close to the critical inclination(8) of 46.4°. The original estimates of Explorer VI lifetime were roughly 200 years without accounting for luni-solar effects; when these perturbations were included, the lifetime dropped to 2 years (!). This lifetime estimate was spectacularly verified by decay before July 1961.

Analytical or semi-analytical work on high-inclination, high-eccentricity resonance phenomena has been almost nonexistent. Classical celestial mechanics essentially ignores this problem since nearly all such orbit problems in the solar system, excepting some asteroids and comets, are blessed with small inclination and eccentricity. Apparently some numerical-analytical work was done by Liouville and Halphen. Musen(9)(10) and Smith(11) have recently revived the methods of Halphen, with modifications by Goriachev of U.S.S.R., and applied them via computer to the Earth satellite problem when the luni-solar influences are (relatively) large. In his initial studies, Musen found that the eccentricity oscillated rather strongly in a large interval, going from large values to zero and increasing again with accompanying inclination changes of as much as 20° (cf., Figs. 2-7). Such pulsating behavior of the eccentricity was also noted by Kozai(12) in his studies of high inclination and eccentricity asteroids perturbed by Jupiter.

Let us return now to the lunar orbiter problem and the question of stationary solutions, or resonance points. From (10) one can readily see that $\dot{\eta}$ is always zero for $\cos g = 0$ or $g = \pm 90°$, for all values of η^2 and β. (To help in relating statements about η^2, β to orbit geometry, one can think of η as "eccentricity" and β as "inclination.") For a resonance condition we must have $\dot{g} = 0$ subject to $\sin g = \pm 1$; this condition can be written as

$$2\left[(1 - \eta^2)(\eta^2 - \beta)\right]^{1/2} \Bigg\{ J_2 a^2 \eta^6 (\eta^2 - 5\beta) + 2x\eta^4 - \frac{5J_4}{32} C_1$$

$$+ 5\left[\frac{J_4}{8} C_2 + x(\beta - \eta^4)\right]\Bigg\} \pm J_3 a\eta^3 C_3 = 0 \qquad (15)$$

where

$$C_1 = \eta^6 + (14\beta + 7)\eta^4 - (63\beta^2 + 126\beta)\eta^2 - 231\beta^2$$

$$C_2 = 5\eta^6 - (56\beta + 7)\eta^4 + (63\beta^2 + 72\beta)\eta^2 - 77\beta^2 \qquad (16)$$

$$C_3 = 4\eta^6 - (35\beta + 5)\eta^5 + (35\beta^2 + 41\beta)\eta^2 - 40\beta^2$$

$$x = \frac{n_E^2 a^7 \eta^4}{\mu}$$

Note that as $e \to 0$, $\eta^2 \to 1$ and the behavior of the zeros of (15) is dominated more and more by the J_3 term. Conclusions concerning the behavior <u>at</u> $e = 0$ cannot be drawn from (10) since the variable g becomes undefined at $e = 0$; characteristics of the zero and near-zero e behavior will be examined later.

The next possibility is that $\cos g \neq 0$ and hence $\dot{\eta} = 0$ implies:

$$\sin g = \frac{-2J_3 a\eta^3(\eta^2 - 5\beta)}{5\left[(1 - \eta^2)(\eta^2 - \beta)\right]^{1/2}\left[J_4(\eta^2 - 7\beta) + 4x\right]} \qquad (17)$$

for $e \neq 0$ and $i \neq 0$, $(\eta^2 = \beta)$. Substituting (17) is $\dot{g} = 0$, we find the resonance condition

$$\frac{5}{J_3^2}\left[J_4(\eta^2 - 7\beta) + 4x\right]^2(1 - \eta^2)(\eta^2 - \beta)\left[J_2 a^2\eta^6(\eta^2 - 5\beta) + 2x\eta^4 - \frac{5J_4}{32} C_1\right]$$

$$- \left[J_4(\eta^2 - 7\beta) + 4x\right]a^2\eta^6(\eta^2 - 5\beta)C_3 + a^2\eta^6(\eta^2 - 5\beta)^2\left[\frac{J_4}{2} C_2\right.$$

$$\left. + 4x(\beta - \eta^4)\right] = 0 \qquad (18)$$

In relations (15) and (18) we have the parameters a and β and ask for the positive roots $\eta^2 \leq 1$ (if any). The implication of (15) and (18) is that we no longer have the critical inclination problem in the usual sense, for now the resonance or stationary solution depends on eccentricity as well as on inclination for a given semi-major axis. In addition, the argument of pericenter is fixed at $g = \pm 90°$ or as determined from (17). In the usual critical inclination problem, as in the case of near-earth satellites, all perturbations except J_2 are second order and hence the stationary solution, to first order, depended only on inclination.

Since the resonance conditions are high-order polynomials in η^2, they must be solved numerically. Prior to discussing the numerical results, however, let us first examine some special cases of the resonance conditions: a) when the eccentricity is close to zero and b) when the semi-major axis is large. As the semi-major axis grows, the Earth influence increases, until at some point the J_n effects may be considered "second order" and hence dropped from the equations. In doing so, the (long-period) equations of motion reduce to

$$\dot{\eta} = -\frac{15}{8L}(1-\eta^2)(1-\beta\eta^{-2})n_E^2a^2 \sin 2g \tag{19}$$

$$\dot{g} = \frac{3n_E^2a^2}{3L\eta^3}\left[(5\beta-\eta^4) - 5(\beta-\eta^4)\cos 2g\right] \tag{20}$$

Thus $\dot{\eta}$ goes to zero at $g = 0°$, $\pm 90°$ and at $\eta^2 = 1, \beta$. For $\eta^2 = 1, \beta$ the eccentricity and inclination are (respectively) zero and we encounter loss of definition of g. For $g = 0°$ the only solution to $\dot{g} = 0$ is for $\eta = 0$ which corresponds to a parabolic orbit of $e = 1$. For the remaining possibilities of $g = \pm 90°$ we find $\dot{g} = 0$ for $\beta = 0.6\eta^4$, a result in agreement with the findings of Kozai,(12) Lidov,(13) Williams and Lorell.(14)

In order to determine the small (and zero) e behavior, introduce the (non-singular) variables

$$A = e\cos g$$
$$B = e\sin g \tag{21}$$

The governing differential equations for these variables are

$$\dot{A} = -\frac{\eta\dot{\eta}}{e}\cos g - B\dot{g}$$

$$\dot{B} = -\frac{\eta\dot{\eta}}{e}\sin g + A\dot{g} \tag{22}$$

Substituting (10) and (11) for $\dot{\eta}$ and $\dot{g}$, simplifying and retaining only terms up to e^2, we find

$$\dot{A} = \frac{3\mu_☾}{4L}\left\{\frac{J_3}{2a^4}(1-\beta)^{1/2}(1-5\beta) + B\left[\frac{n_E^2a^2}{\mu_☾}(3-5\beta) - \frac{J_2}{a^3}(1-5\beta)\right.\right.$$
$$\left.\left. + \frac{5J_4}{16a^5}(8-88\beta-119\beta^2)\right] + \frac{J_3B^2}{2a^4(1-\beta)^{1/2}}(5-41\beta+40\beta^2)\right.$$

$$+ \left. \frac{J_3(A^2 + B^2)}{4a^4(1 - \beta)^{1/2}} (4 - 35\beta + 35\beta^2) \right\} \tag{23}$$

$$\dot{B} = \frac{3\mu_{\mathbb{C}}}{4L} \left\{ \frac{J_4}{a^3} (1 - 5\beta) + \frac{2n_E^2 a^2}{\mu_{\mathbb{C}}} - \frac{5J_4}{16a^5} (4 - 56\beta - 147\beta^2) \right.$$

$$\left. - \frac{J_3 B}{2a^4(1 - \beta)^{1/2}} (5 - 41\beta + 40\beta^2) \right\} A \tag{24}$$

For $e = 0$ the only possible stationary solutions are for $\beta = 1$ and $\beta = 0.2$ corresponding to inclination of $i = 0°$ and $i = 63.4°$ respectively. Note that for small e, say $e \leq 0.1$, and small $\beta(\beta \leq 0.5)$ we can integrate (23) and (24) by successive approximation with J_3e^2 considered to be "second order." The integration technique parallels that of Breakwell and Hensley[8] except that in the present case we no longer have the symmetry of coefficients and variables.

In the general case, when the eccentricity may take any (allowable) value we determine the stationary solutions from (15) and (18) via computer. This was done for the two sets of J_n values quoted in Section III and the results shown in Figs. 9 and 10 for the Langley and U.S.S.R. values respectively. Two major divisions exist in these figures; the solutions of (15) for $g = \pm 90°$ and of (18) with g determined from (17), the former appearing in the upper left, the latter in the left center of the figures. The loci of critical inclination and eccentricity for a (representative) range of semi-major axis values are given, terminating at the maximum allowable value of the eccentricity for the semi-major axis in question. For the $g = \pm 90°$ loci the $+90°$ loci are shown in dotted lines, the $-90°$ loci as solid lines. As noted earlier, these tend to the two zero eccentricity solutions of $i = 0°$ and $i = 63.4°$ as $e \to 0$. Identified by its equation, $\cos^2 i = 0.6(1 - e^2)$, is the Earth-only curve which is the limiting locus for high-a orbits; the locus for $a = 5$ is indistinguishable for non-zero e from the limiting curve on the scale shown. Superimposed on the loci for $g \neq \pm 90°$ in the upper left corner are contours of constant g values. The upper left corner region ceases to exist for $a > 2$ since then the stationary solutions move "under the lunar surface," i.e., they exist for eccentricities greater than the maximum allowable e.

An interesting feature can be observed in the evolution of the $\pm 90°$ loci as the semi-major axis is decreased from, say 5. On the limiting curve the loci for $\pm 90°$ coincide. As one decreases the semi-major axis the plus and minus curves separate due to the increasing influence of the odd harmonic J_3 (multiplier of J_3 term is $\sin g$). The critical inclination for a given eccentricity is greater for $g = +90°$ than for $g = -90°$ if $J_3 > 0$ and vice versa if $J_3 < 0$ (cf. Figs. 9, 10). Furthermore, in the case of $J_3 > 0$, the maximum inclination attainable for $g = -90°$ by decreasing a peaks out at about $a = 1.4$, then decreases as $a \to 1$. On the other hand, for $g = +90°$, the maximum inclination increases monotonically and goes to $63.4°$ as $a \to 1$.

The situation is modified when the J_3 and J_4 coefficients change sign. (From the present results it is difficult to estimate the influence of the slight J_2 value decrease, but presumably the majority of the changes are due to the (significantly) different values of J_3 and J_4.) The $\pm 90^\circ$ loci still tend to separate with decreasing semi-major axis but now the curves tend to draw together again as a decreases below roughly 1.7, say. This is due to the proximity of the low-a curves to the inclination 63.4° for which the J_3 influence is minimum. The inclination for a stationary solution at a given semi-major axis and eccentricity is larger for the coefficients given by the Luna 10 flight than for the Langley coefficients for all values of a off the limiting curve.

The last point to consider is the nature of the stationary solutions: Are they stable or unstable? Stability in this context means bounded variations of both η^2 and g for motion near a stationary point; furthermore, the variation of η^2 must be such as not to exceed η^2_{crit}. This question can be answered from analysis of the Hamiltonian and its partial derivatives at a stationary point. Since variation of g and η^2 must be bounded near a maximum or a minimum of the Hamiltonian, the stationary points corresponding to the minimum and maximum are stable. The saddle points of the Hamiltonian will give the unstable stationary points. These conditions are easily checked on the computer at the time of computation of the stationary solutions from the value of the determinant

$$\Delta = \begin{vmatrix} \dfrac{\partial^2 \mathcal{F}^*}{\partial g^2} & \dfrac{\partial^2 \mathcal{F}^*}{\partial g \partial G} \\ \dfrac{\partial^2 \mathcal{F}^*}{\partial G \partial g} & \dfrac{\partial^2 \mathcal{F}^*}{\partial G^2} \end{vmatrix} \tag{25}$$

From the canonical nature of the variables g, G, we have Δ as

$$\Delta = \begin{vmatrix} \dfrac{\partial \dot{G}}{\partial g} & \dfrac{\partial \dot{G}}{\partial G} \\ \dfrac{\partial \dot{G}}{\partial G} & -\dfrac{\partial \dot{g}}{\partial G} \end{vmatrix} \tag{26}$$

so that stable stationary solutions are characterized by

$$\Delta \geq 0 \quad \text{and} \quad \frac{\partial \dot{G}}{\partial g} > 0 \quad \text{or} \quad \frac{\partial \dot{G}}{\partial g} < 0 \tag{27}$$

and unstable stationary solutions by

$$\Delta < 0 \tag{28}$$

All of the stationary solutions of Figs. 9 and 10 were found to be stable, corresponding to maxima of the Hamiltonian for $g \neq \pm 90^\circ$ and minima for $g = \pm 90^\circ$. The "saddle points" of the Hamiltonian, corresponding to stationary solutions where equi-energy contours in the η^2, g plane intersect, presumably exist only for $\eta^2 < \eta^2_{crit}$.

Conclusions

In view of the results of the preceding section some additional comments on the η^2 - g contours of Figs. 2-7 are in order. To aid in understanding the evolution of the curves one can plot the locus of stationary η^2 versus β for a constant semi-major axis; this is done in Figs. 11 and 12 for $a = 1.7$ and $a = 3.0$ respectively for the Langley set of J_n's. The figures pertinent to this dicussion then are 2, 3, 4 and 11, 12.

As β increases from zero, a stationary point develops at the appropriate value of g (Fig. 9) and moves "downward" in the decreasing η^2 direction, disappearing under the lunar surface. Increasing β still further, we encounter the $g = +90°$ stationary point "rising" from below η^2_{crit} soon joined by the $g = -90°$ stationary solution. Further increase in β drives both $+90°$ and $-90°$ stationary solutions toward $\eta^2 = 1$, and purely circulating (and, for the most part, surviving) orbits remain. For $a > 2$ we have no non $90°$ stationary solutions but otherwise the evolution is similar.

For high-a orbits the influence of the J_n's is slight; hence the $a = 3$ picture for the U.S.S.R. J_n's is not presented, being only a slight variation of Fig. 4. The primary difference for the $a = 1.7$ case with the U.S.S.R. J_n's is the "evening" of the $\pm 90°$ stationary solutions on the η^2 axis, implying that the critical inclinations are somewhat closer together for $g = \pm 90°$ than in the case of the Langley J_n's.

In Figs. 11 and 12 for η^2 very close to unity one seems to pick up an additional stationary η^2 solution at a given β value for $|g| \neq 90°$ and for $g = +90°$. The existence of such double solutions could not be verified from the equi-energy contour program due to numerical resolution difficulties. It seems that it would be more advantageous to investigate these regimes via the low and zero e equations. Although the rapid falling-off in inclination at low e (Figs. 9 and 10) was verified for the $g = \pm 90°$ stationary solutions from the small-e equations in A and B, more work is needed in this area. As pointed out earlier, probably more progress could be made for the restricted eccentricity problem through approximate integrations of the $\dot{A}$ and $\dot{B}$ equations and thus the behavior very close to $\eta^2 = 1$ clarified.

Finally, it can be verified that the behavior for semi-major axis in the range $1 < a < 2$ is similar to that presented for $a = 1.7$, with variations only as to the location of stationary points with β and η^2; for $a \geq 2$ the results are similar to those presented for $a = 3.0$.

References

Brouwer, D.: "Solution of the Problem of Artitical Satellite Theory Without Drag," _Astronomical Journal_, 64, 1959.

Kozai, Y.: "Motion of a Lunar Orbiter," _Journal of the Astronomical Society of Japan_, 15, 1963.

Garfinkel, B.: "The Disturbing Function for an Artifical Satellite," _Astronomical Journal_, 70, 1965.

Vagners, J.: "Some Resonant and Non-Resonant Perturbations of Earth and Lunar Orbiters," Stanford University Department of Aeronautics and Astronautics report SUDAAR 317, Stanford University, Stanford, California.

Oesterwinter, C.: "The Motion of a Lunar Satellite," Ph.D. Dissertation Presented to Yale University, 1965.

Giacaglia, G.E.O. et al.: "The Motion of a Satellite of the Moon," Goddard Space Flight Center Report X-547-65-218, Greenbelt, Maryland, June 1965.

Akim, E.L.: "Determination of the Gravitational Field of the Moon by the Motion of the AMS LUNA-10," Translated from Doklady A.N. SSR., Mechanika Tom 170, no. 4 Izdatel'stoo "NAUKA," 1966, by A. L. Brichant, NASA (NASA No. ST-CM-LPS-10532).

Breakwell, J.I. and D. Hensley: "An Investigation of High Eccentricity Orbits about Mars," in First Compilation of Papers on Trajectory Analysis and Guidance Theory, NASA SP-141, 1967.

Musen, P.: "On the Long Period Lunar and Solar Effects in the Motion of an Artificial Satellite," Journal of Geophysical Research, 66, 1961.

Musen, P.: "Discussion of Halphen's Method for Secular Perturbations and its Application to the Determination of Long Range Effects in the Motion of Celestial Bodies" Part 1. NASA Technical Report R-194, 1963.

Smith, A.J.: "A Discussion of Halphen's Method for Secular Perturbations and its Application to the Determination of Long Range Effects in the Motion of Celestial Bodies" Part 2. NASA Technical Report R-194, 1964.

Kozai, Y.: "Secular Perturbations of Asteriods with High Inclination and Eccentricity," Astronomical Journal, 67, 1962.

Lidov, M.L.: "Evolution of the Orbits of Artificial Satellites of Planets as Affected by Gravitational Perturbations From External Bodies" Russian Supplement, AIAA Journal vol. 1, August 1963.

Williams, R.R. and Lorell, J.: " The Theory of Long-Term Behavior of Artificial Satellite Orbits due to Third-Body Perturbations" Jet Propulsion Laboratories Technical Report 32-916 February 1966.

THE EFFECT OF THREE-DIMENSIONAL, NON-LINEAR RESONANCES ON THE MOTION OF A PARTICLE NEAR THE EARTH-MOON EQUILATERAL LIBRATION POINTS

By Hans B. Schechter
Department of Aeronautics and Astronautics
Stanford University
Stanford, California

NASA Research Grant NsG 133-61

ABSTRACT

In the uniformly rotating reference frame of the restricted 3-body problem (in which Earth and Moon occupy fixed positions on the abscissa), the equilateral libration points L and L are known to be points of equilibrium. A particle placed at rest at one of these points will remain at rest for all times. According to linear theory, for very small disturbances from equilibrium the particle will tend to move along bounded trajectories in the immediate vicinity of these points.

When the force field near L and L is not assumed to be linear, and in addition other perturbing effects are included, the particle's motion might be excited sufficiently and lead to unstable divergent trajectories.

This report presents the results of an analytic study of the 3-dimensional stability of motion of a particle near L_4 in a nonlinear Earth-Moon force field, upon which is superimposed a linear solar gravitational field distribution. In particular, the long period features of the particle's motion are studied, which stem from the excitation at or close to the particle's natural frequencies, and are introduced by the presence of resonance terms in the internal (Earth and Moon) and external (solar) force fields.

The results show that in the presence of the internal nonlinearities the stability of motion predicted by the linear theory is valid for only a very restricted region of initial displacement and velocity disturbances. Disturbances outside this region would lead to divergence of the solution. The nonlinear coupling of the out-of-plane terms with the in-plane terms was found to be of minor importance and did not contribute to an appreciable transfer of energy from one mode of motion to the other.

The inclusion of the external force terms was found to admit some equilibrium solutions of the variational equations. Of those, the one

stable equilibrium solution found was characterized by a coplanar elliptic particle orbit around L_4 which had its major axis (of magnitude roughly 120,000 mi) oriented at right angles to the line joining Earth to L_4. This orbit was traversed in a clockwise sense at mean angular rate equal to that of the Sun, as seen in the rotating coordinate frame, and very close to the particle's faster coplanar natural frequency. The particle's motion thereby became synchronized with that of the Sun.

LIST OF SYMBOLS

C_J	=	Jacobi constant introduced in Eq. (43)
D	=	mean Earth-Moon distance defined by Eq. (8)
D_1	=	integration constant introduced in Eq. (65)
$e \simeq .055$	=	eccentricity of lunar orbit
f, η	=	functions of ξ defined in Eqs. (69) and (70)
$f_x, f_y \cdots$	=	forcing functions introduced in Eqs. (37)
G	=	Universal gravitational constant
$G(\beta,\alpha'), S(\beta,\alpha')$	=	generating functions introduced in Eq. (45)
H	=	Hamiltonian = $H^{(o)} + H'$
H'	=	contains the higher order nonlinearities
$\tilde{\bar{H}}_3, \tilde{\bar{H}}_4, \bar{H}_4$	=	slowly varying Hamiltonians used in Eq. (48)
H_r, H_{p_r}	=	partial derivatives of H as defined in Eqs. (22)
H_s	=	Hamiltonian resulting from solar effects
$H^{(o)}$	=	contains the linear and quadratic terms
$H^{(o)}_{\alpha_i}, H^{(o)}_{\beta_i}$	=	partial derivatives of $H^{(o)}$ as defined in Eqs. (41)
i	=	inclination of E-M plane with ecliptic
J	=	transformation matrix defined in Eq. (32)
$J(\beta',\alpha',t)$	=	$J_1 + J_2 + J_3$ = generating function introduced in Section IX.
K'	=	Hamiltonian containing only secular and slowly varying terms
K^*	=	time independent Hamiltonian = $K_i^* + K_e^*$
K_{i2}^*	=	coplanar part of K^* defined in Eq. (63)
L	=	Lagrangian

LIST OF SYMBOLS

m = n_s/n = .074801 = dimensionless Earth orbital angular velocity

m_i = mass of ith body

n = mean angular velocity of E-M system

n_s = mean angular velocity of Earth around Sun

P_x, P_y, P_z; p_x, p_y, p_z = momenta defined by Eqs. (19)

Q, P = normal canonical coordinates introduced via Eq. (30)

$\bar{R}$ = displacement vector in inertial space

r_{ij} = distance between masses

$\bar{r}$ = position vector measured from L_4

$\bar{r}_{1L}$ = position vector from Earth to point L_4

T = kinetic energy

$T_\alpha = 2\pi/\omega_\alpha$ = period of slow oscillation in α_1^* and α_2^* obtained from Eq. (84)

V = potential energy

W = normal out-of-plane solar acceleration at Moon's position, introduced in Eqs. (121)

$\bar{x}, \bar{y}, \bar{z}$ = solutions to homogeneous linear equations

$\tilde{x}, \tilde{y}, \tilde{z}$ = forced response of linearized system

x_s, y_s, z_s = solar coordinates in xyz system, defined in Eq. (17)

α', β' = set of "slowly varying integration constants"

α^*, β^* = set of variables canonical with respect to K^*

LIST OF SYMBOLS

α_i, β_i	=	set of integration constants
$\beta_i^{\neq}$	=	$\left[t - (-1)^i \beta_i\right]$ where $(i = 1,2,3)$
$\Delta^* = \beta_1^* + \beta_2^*$	=	angular variable
Δ_{13}	=	angular variable defined in Eq. (51)
$\Delta\omega_1, \Delta\omega_2$	=	frequency shifts in the coplanar natural frequencies, defined in Eq. (79)
$\xi, \phi, \eta_o, \Omega, \upsilon'$	=	angular variables introduced in Eqs. (15), (16), and (17) and defined in Appendix B
$\epsilon_{12}, \epsilon_{13}, \epsilon_{\xi 1}$	=	detuning frequencies
η	=	constant defined in connection with Eq. (26)
μ	=	dimensionless quantity defined by Eq. (12) = $m_M/(m_E + m_M) \cong 1/82.45$
$\xi = \alpha_2^*/D_1$	=	variable introduced by Eq. (67)
ρ, υ	=	perturbation quantities defined by Eqs. (15) and (16)
Φ_o	=	6 × 6 matrix defined by Eq. (24)
$\dot{\Omega}, i$	=	angular velocities of line of nodes, and of E-M plane inclination, respectively
ϖ	=	angular velocity vector of xyz system
$\omega_1, \omega_2, \omega_3$	=	eigenvalues of the linear homogeneous set of differential equations (i.e., the natural frequencies about L_4)
ω_M	=	angular velocity of a hypothetical isolated E-M system (i.e., no solar perturbations present)

LIST OF SYMBOLS

$(\ )^T$ = denotes transpose of ( )

$(\dot{\ })$ = denotes total time derivative

I. INTRODUCTION

The subject of the Earth-Moon libration points has aroused in recent years the curiosity and interest of a great many researchers in the field of celestial mechanics and analytical dynamics. This renewed interest by modern day investigators in this classical problem has been stimulated by the recent telescopic sightings by K. Kordylewski[1,2] of two faint cloud-like objects or shapes in the vicinity of the L_4 and L_5 Earth-Moon libration points. These findings have led to a great amount of speculation regarding the origin and stability of motion of such clouds, believed by many to be composed of minute dust particles.

Although a number of more recent naked eye sightings from high flying aircraft have since been reported by a few investigators in this country, the issue of the existence or nonexistence of these libration dust clouds has not yet been resolved to everyone's satisfaction by any of the current studies, and is still the subject of debate between proponents and detractors of this hypothesis. While the definitive answer to this question might not be obtained until concrete evidence and data will be gathered near these points from a space vehicle, the quest so far has not been all in vain. In the process a great many areas for further research of both a theoretical and a practical, mission oriented, nature have been exposed and tackled, which will keep many researchers busy for quite a while.

In the present dissertation we shall not attempt to shed new light on the question of the existence of dust clouds, but shall confine instead our attention to the study of the interesting underlying theoretical problem in nonlinear analytical dynamics of a particle. This particle may be associated, if one desires to do so, with the center of mass of a hypothetical dust cloud. It should be pointed out however that the uncritical application of some of the results and conclusions of the present study to the dust cloud problem might lead to misleading conclusions, since such important destabilizing effects as solar radiation pressure and particle collisions have not been considered here.

II. LIBRATION POINT GEOMETRY

Some of the geometrical features of libration points are briefly indicated below for the purpose of orientation.

The five libration points (also known as Lagrangian points) of the classical restricted 3-body problem (i.e., Sun is neglected, and Earth and Moon revolve in circular orbits about their common center of mass) are indicated in Fig. 1. They are points of equilibrium in the coordinate frame XYZ, rotating around the Z axis with the mean angular velocity n of the Earth-Moon system, in the sense that no net accelerations are experienced by particles at rest at these points.

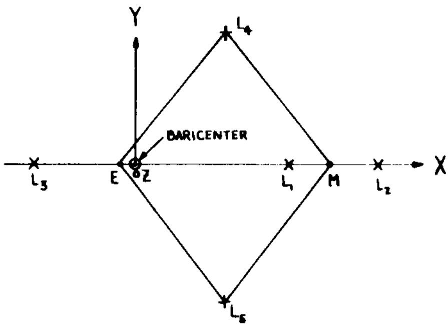


Fig. 1: Libration points of the restricted 3-body problem.

By means of linear small perturbation analysis the collinear points L_1, L_2, L_3 were found to be unstable to small initial disturbances, while the equilateral points L_4 and L_5 were found to be points of stable equilibrium around which small amplitude conditionally periodic (i.e., in this case doubly periodic but not necessarily simply periodic) motions resulted for small initial disturbances.

The more realistic physical model used in the present analysis is shown in Fig. 2. The Sun, lunar orbital eccentricity e ($\approx$ .055) and inclination i of the Earth-Moon plane with the ecliptic ($i \approx 5^o$) are included. The Earth is assumed to move in a circular orbit around the Sun.

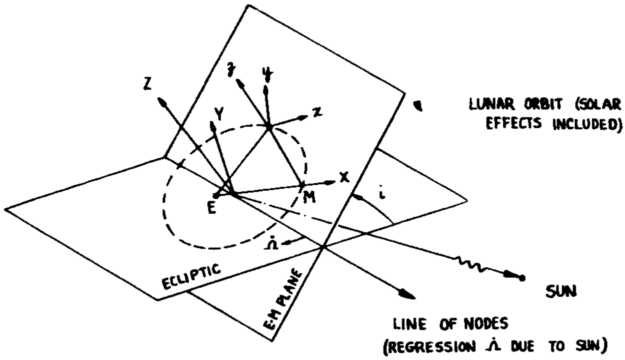


Fig. 2: Three dimensional geometry of the 4-body problem.

III. BRIEF REVIEW OF PAST WORK ON THE SUBJECT

Most of the basic work on the restricted 3-body problem stems back to some of the classical studies in analytical dynamics of Lagrange, Jacobi, Poincaré, etc. which are discussed in most of the standard textbooks on Celestial Mechanics. Some of the main features and results are briefly summarized in the following sections.

More recent analytic work on the 3-body problem concerned itself with such questions as the existence of periodic orbits both in the vicinity of the libration points, as well as periodic orbits which fill the whole Earth-Moon space and possibly loop a number of times around both primary bodies.

Studies which included the solar force field are of a more recent vintage and are predominantly of a numerical nature, in that they tackle the problem by direct integration of the full set of differential equations of motion for various periods of time t, and usually for a very restricted set of initial conditions[3-5] (i.e., zero particle displacements and velocities, and collinear position of the major bodies in the order Earth-Moon-Sun). The application of Hamiltonian techniques to the 2-dimensional libration point problem was suggested in an analytic study by Breakwell and Pringle.[6] These techniques are extended in the present thesis to the 3-dimensional problem which also includes the effects of lunar orbital eccentricity.

1. THE CLASSICAL RESTRICTED 3-BODY PROBLEM: PAST RESULTS AND THEIR LIMITATIONS

Some of the basic results of the 3-body theory, as related to the libration points, and some of the questions left unanswered by the theory are mentioned in A and B, respectively.

A. 1. The existence of the five Lagrangian equilibrium points shown in Fig. 1 was discovered.

2. The stability of motion near these points was investigated by linearizing the equations of motion near these points.

3. For small deviations from equilibrium the coplanar homogeneous set of equations (Eqs. (25) with $\rho = \upsilon = m = 0$) in the xy plane, which becomes uncoupled from the z equation, was shown to give rise to a doubly periodic solution with the eigenvalues $\omega_1 \simeq .955$ and $\omega_2 \simeq .298$ (these frequencies were nondimensionalized with respect to the mean Earth-Moon angular velocity $n \simeq .23$ rad/day). The uncoupled, out of plane, linear equation in the z direction possesses a simple harmonic solution with eigenvalue $\omega_3 = 1$. (The reason for a period of 1 lunar month in the z motion is easy to explain physically if we consider the limiting case of a vanishingly small lunar gravitational force field. In that case the small particle at L_4 follows a near circular planar 2-body orbit around the Earth at the lunar distance, which crosses the Earth-Moon plane twice for each complete particle revolution, thus leading to an orbital period of 1 lunar month, which is also the same as the period of the projected simple harmonic oscillator in the z direction.)

4. A first, and only, integral constant of the motion was found to exist. This so-called Jacobi constant C_J corresponds to our scleronomic (i.e., time independent) Hamiltonian H, and consists of the combination $E - nh_z = \text{constant} = - C_J = H$, where E is the particle's total energy (i.e., kinetic and potential) in a nonrotating baricenter centered coordinate frame, h_z is its angular momentum in the Z direction, and n is the mean angular velocity of the Earth-Moon axis.

B. Some difficulties are encountered if one tries to extend the stability conclusions obtained from linear analysis to predict the behavior of the complete nonlinear system. The main reasons are indicated below.

1. The near commensurability of the eigenvalues $\omega_1 \simeq 3\omega_2$ leads to an internal near resonance with a detuning $\epsilon_{12} = \omega_1 - 3\omega_2 \simeq .954593-3\cdot.297912 = \simeq .06086$. This causes poor convergence of the usual perturbation solutions by means of which one attempts to evaluate the effects of higher order terms, by substituting back the homogeneous solutions into the nonlinear driving terms. Some of them give rise to combination frequencies which are nearly resonant with the natural frequencies of the linear equations, and thus lead to small divisors in the next approximation.

2. The Hamiltonian H is not definite near L_4 or L_5 (positive or negative). This sign indefiniteness has a bearing on the nature of the stability of adjacent motions, as is briefly indicated below.

If we suitably recombine the terms in H of A(4) above we can come up with an equivalent relation for the Hamiltonian $H = \frac{1}{2} v^2 + V_{eff}$, where $v^2 = \dot{x}^2 + \dot{y}^2 + \dot{z}^2$ and V_{eff} represents an effective potential energy $V_{eff} = -\frac{1}{2} \omega^2 (x^2 + y^2) - \mu_1/r_1 - \mu_2/r_2$. The first term in H thus corresponds to the kinetic energy, as measured in the rotating frame, while the last two terms in V_{eff} represent the usual gravitational potential energy V. In this new form H can be interpreted as being in the nature of an energy integral of the motion. The nature of the stability near $L_{4,5}$ can thus be deduced from the shape of the surfaces V_{eff} = constant in that region. It turns out that near the equilateral points the planar part of V_{eff} has the shape of a "potential hill" rather than the "trough" which is required for stability.

This circumstance raises a question concerning the applicability of the linear-theory stability analysis to the complete nonlinear system, i.e., whether the nonlinear system would exhibit the same kind of stability as predicted by the linear equations for given initial conditions. One may remark at this point, on the basis of work to be presented later, that the answer is yes in a rather small neighborhood of L_4. The nonlinear system will however exhibit instability for certain ranges of initial conditions.

It is also appropriate to remark here that the stability of motion exhibited by the linear system near L_4 and L_5 in the presence of a potential energy "hill" is brought about by the presence of gyroscopic terms in the linear equation (due to the Corioli's force $2(\bar{n} \times \dot{\bar{r}})$ which arise in the rotating frame). When further nonlinear and external effects are included, it is possible for additional energy to be transferred into the system with the result that initially small oscillations may grow in the course of time.

It is interesting to mention that a Taylor series expansion of V_{eff} near L_4 shows the equipotential curves to be extremely elongated ellipses of fineness ratio roughly 1:10 oriented at right angles to the line from barycenter to L_4. The potential field thus falls off quite slowly as we move in a direction perpendicular to the Earth - L_4 line.

3. Another internal resonance occurs because of nonlinear coupling of the z and xy solutions, and the near commensurability of the eigenvalues $\omega_1 \cong \omega_3$, with the resulting detuning $\epsilon_{13} \cong .0454$. This resonance leads again to poor convergence of perturbation type solutions.

4. Although not actually a part of the classical 3-body problem, it might perhaps not be inappropriate to mention at this point also the presence of a third important resonance of an external nature caused by the Sun's perturbative action on a nominally circular lunar orbit, which is an important factor in the subsequent analysis. This indirect solar perturbation leads to a detuning $\epsilon_{\xi 1} = 2[\omega_1 - (1 - m)] = 2[.95459-.92520] \simeq .05878$.

5. The additional complications of resonances introduced by the inclusion of lunar eccentricity terms will be taken up later.

2. NUMERICAL APPROACHES (SOLAR EFFECT INCLUDED)

Straightforward integration of the complete set of differential equations, for zero initial conditions, gives rise to particle trajectories, a typical xy projection of which looks roughly like the one shown in Fig. 3 (taken from Ref. 4).

Figure 4 presents schematically another plot due to Feldt and Shulman[5] of total particle displacement d with time t for an integration time period of 5000 days. Initial conditions were the same as those in Fig. 3.

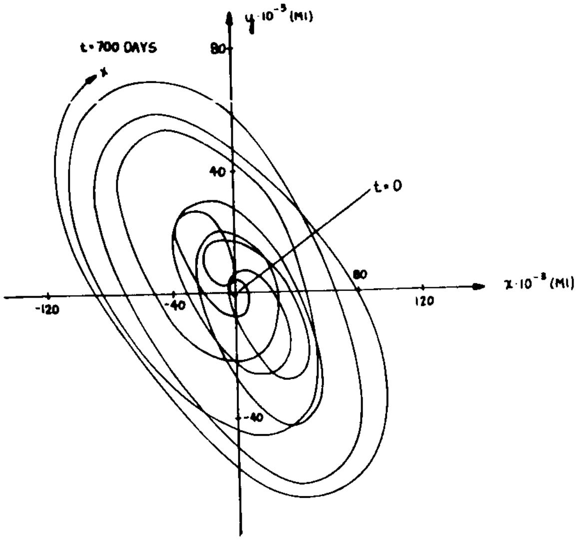


Fig. 3: Typical particle trajectory in xy plane near L_4 (t = 700 days)

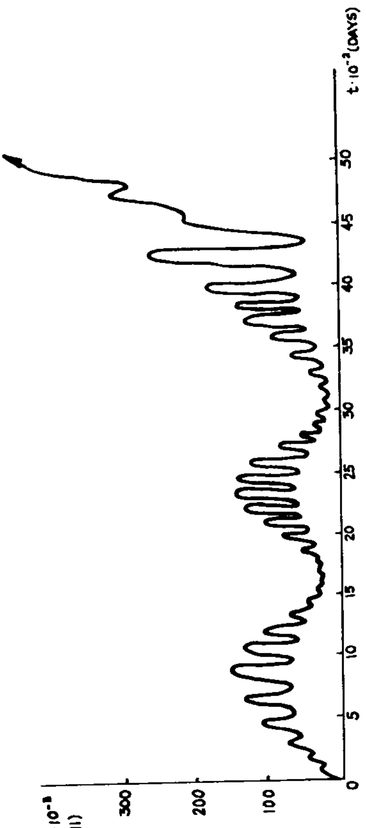


Fig. 4: Displacement-time history

IV. SOME CONCLUSIONS REGARDING PRIOR STATUS OF THE PROBLEM

The following conclusions summarize some of the points which were raised in Sections III(A) and III(B):

1. The past analytical efforts do not resolve in a satisfactory manner the question of boundedness of motion near the equilateral libration points of the Earth-Moon system, with or even without the inclusion of the perturbative effect of the Sun.

2. The numerical results available to date are rather limited in that they were generated only for restricted sets of initial conditions and initial Earth-Moon-Sun configurations. Consequently they do not shed much further light on the question of the possible existence of domains of initial conditions and configurations which allow small amplitude, bounded motions to take place for long time periods.

3. In view of the multiplicity of possible starting conditions and configurations, it is quite clear that a purely numerical search for such initial conditions would be both costly as well as of questionable success, and thus not very attractive.

4. The necessity and usefulness for further analytical groundwork on this problem seems to be clearly indicated.

The above brief rundown will hopefully help to bring into better perspective the difficulties as well as the motivations underlying the present investigation.

V. THE LAGRANGIAN L FOR A PARTICLE NEAR L_4

We shall desire the expression for the Lagrangian of a particle near the L_4 libration point, in the rotating xyz frame centered at L_4, and having its xy plane coincide with the fundamental Earth-Moon orbital plane. To this end it is convenient to start out with an inertial reference frame x_I, y_I, z_I in which the positions of Earth, Moon, Sun and particle P are designated by the numbers 1, 2, 3, and 4, respectively, and by the position vectors $\bar{R}_i$ $(i = 1,\cdots 4)$. The kinetic energy T_I and potential energy V_I of all the masses are then

$$T_I = \frac{1}{2} \sum_{i=1}^{4} m_i \dot{\bar{R}}_i \cdot \dot{\bar{R}}_i$$

$$V_I = -\frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^{4} \frac{G m_i m_j}{|\bar{R}_i - \bar{R}_j|} = -\frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^{4} \frac{G m_i m_j}{r_{ij}} \tag{1}$$

We switch first to an Earth centered rotating coordinate system X_e, Y_e, Z_e with the X_e axis pointing in the direction of the instantaneous position of the Moon (we neglect here the 3000 mi separation of barycenter from the center of the Earth). For a particle of unit mass at point 4 we then have

$$T = \frac{1}{2}\left(\dot{\bar{R}}_1 + \dot{\bar{r}}_{14}\right) \cdot \left(\dot{\bar{R}}_1 + \dot{\bar{r}}_{14}\right)$$

$$V = -\frac{\mu_1}{r_{14}} - \frac{\mu_2}{r_{24}} - \frac{\mu_3}{r_{34}} \tag{3}$$

$$L = \frac{1}{2}\left[\dot{\bar{r}}_{14} \cdot \dot{\bar{r}}_{14} + 2\dot{\bar{R}}_1 \cdot \dot{\bar{r}}_{14}\right] + \frac{\mu_1}{r_{14}} + \frac{\mu_2}{r_{24}} + \frac{\mu_3}{r_{34}} + \frac{1}{2}\dot{\bar{R}}_1 \cdot \dot{\bar{R}}_1$$

where

$$\mu_i = Gm_i$$
$$i = 1,2,3$$

The last term in L is independent of particle position and velocity and can be dropped. This follows from our assumption that the particle does not affect the motion of the primary bodies. It is also convenient to remove from L the explicit presence of the Earth's inertial velocity $\dot{\bar{R}}_1$. This can be done via Lagrange's equation

$$\frac{d}{dt}\left[\frac{\partial L}{\partial \dot{\bar{r}}_{14}}\right] - \frac{\partial L}{\partial \bar{r}_{14}} = 0 \tag{4}$$

and the Earth's equation of motion in inertial space

$$\ddot{\bar{R}}_1 = \frac{\mu_2}{r_{12}^3}\,\bar{r}_{12} + \frac{\mu_3}{r_{13}^3}\,\bar{r}_{13} \tag{5}$$

Since $\ddot{\bar{R}}_1 \neq \ddot{\bar{R}}_1\,(\bar{r}_{14})$, one can replace Eq. (5) by the equivalent relation

$$\ddot{\bar{R}}_1 = \frac{\partial}{\partial \bar{r}_{14}}\left[\frac{\mu_2}{r_{12}^3}\,\bar{r}_{12}\cdot\bar{r}_{14} + \frac{\mu_3}{r_{13}^3}\,\bar{r}_{13}\cdot\bar{r}_{14}\right] \tag{6}$$

After substituting Eq. (6) into (4) one can extract from it the expression for L shown in Eq. (7):

$$L = \frac{1}{2}\,\dot{\bar{r}}_{14}\cdot\dot{\bar{r}}_{14} + \frac{\mu_1}{r_{14}} + \mu_2\left(\frac{1}{r_{24}} - \frac{\bar{r}_{12}\cdot\bar{r}_{14}}{r_{12}^3}\right) + \mu_3\left(\frac{1}{r_{34}} - \frac{\bar{r}_{13}\cdot\bar{r}_{14}}{r_{13}^3}\right) \tag{7}$$

The last (solar) term in (7) can be further simplified if we replace it with the solar potential energy gradient evaluated at the position of the Earth, as shown in Appendix A. This neglects terms of magnitude $(r_{14}/r_{13})^3 \cong 1.5 \times 10^{-8}$, which is quite satisfactory in the present case, and leads to the expression

$$L = \frac{1}{2}\dot{\bar{r}}_{14} \cdot \dot{\bar{r}}_{14} + \frac{\mu_1}{r_{14}}$$

$$+ \mu_2\left(\frac{1}{r_{24}} - \frac{\bar{r}_{12} \cdot \bar{r}_{14}}{r_{12}^3}\right) + \frac{\mu_3}{r_{13}^3}\left[\frac{3}{2}\left(\frac{\bar{r}_{13} \cdot \bar{r}_{14}}{r_{13}}\right)^2 - \frac{1}{2}\bar{r}_{14} \cdot \bar{r}_{14}\right] \quad (8)$$

Expression (8) is still not in the desired final form of a Taylor series expansion around L_4. Before we carry out the expansion it is convenient to nondimensionalize everything, as indicated in the next section.

VI. NONDIMENSIONALIZATION AND EXPANSION AROUND L_4

1. NONDIMENSIONALIZATION

The nondimensionalization is most conveniently carried out by choosing the reference frequency n and length D defined by

$$n = \sqrt{\frac{\mu_1 + \mu_2}{D^3}} = < \omega_M + \dot{\Omega} \cos i > = \text{mean angular velocity of E-M axis } X_e \cong .23 \text{ rad/day} \tag{9}$$

$$D = < r_{12} > = \text{mean E-M distance} \cong 2.4 \times 10^5 \text{ mi} \tag{10}$$

It should be pointed out that the only physical quantity which can be measured with any degree of accuracy is n, so that the reference length D is actually a computed, rather than a natural quantity, and is defined by Eq. (9). The averaging of r_{12} in Eq. (10) must therefore be interpreted in the light of the more basic definition (9).

ω_M denotes the mean angular velocity of an isolated Earth-Moon system (no solar perturbations present), and $\dot{\Omega}$ and i are indicated in Fig. 2.

Two basic dimensionless quantities which will appear often in our equations are

$$m = \frac{n_s}{n} = \sqrt{\frac{\mu_3}{r_{13}^3} \cdot \frac{D^3}{\mu_1 + \mu_2}} \cong .074801 \tag{11}$$

and

$$\mu = \frac{\mu_2}{\mu_1 + \mu_2} \cong \frac{1}{82.45} \tag{12}$$

where

n_s = angular velocity of the Earth around the Sun

From now on all lengths, velocities and times will be treated as dimensionless quantities, but we shall retain their old symbols.

2. EXPANSION AROUND L_4

Just as n was the basic quantity selected in the nondimensionalization of the equations, we shall select m as the basic quantity, or yardstick, which defines order of magnitude. We shall denote by o(m) a quantity of first order of smallness, $o(m^2)$ of second order, etc...

The Lagrangian L of Eq. (8) can be written in terms of displacements and velocities measured in the L_4 centered xyz frame by writing the dimensionless vector relations

$$\begin{aligned} \bar{r}_{14} &= \bar{r}_{1L} + \bar{r} \\ \dot{\bar{r}}_{14} &= \dot{\bar{r}}_{1L} + \dot{\bar{r}} + \bar{\omega} \times \bar{r} \end{aligned} \tag{13}$$

where

$$\bar{r} = x\bar{i}_x + y\bar{i}_y + z\bar{i}_z$$

$|\bar{r}_{1L}| = 1 + \rho(t) = |\bar{r}_{12}|$ = instantaneous displacement of libration point L_4 from the Earth

$$\dot{\bar{r}}_{1L} = |\dot{\bar{r}}_{1L}|\left(\frac{1}{2}\,\bar{i}_x + \frac{\sqrt{3}}{2}\,\bar{i}_y\right) + \bar{\omega} \times \bar{r}_{1L}$$

and for the total angular velocity $\bar{\omega}$ of the xyz frame in inertial space

$$\bar{\omega} = \frac{n}{n}\,\bar{i}_z + \bar{\upsilon}(t) = \bar{i}_z + \bar{\upsilon}(t) \tag{14}$$

$\rho(t)$ and $\upsilon(t)$ are the perturbations of the E-M distance, and angular velocity caused by solar and eccentricity effects, and are provided by classical lunar theory.[7,8]

$$\rho(t) = - .0079 \cos 2\xi - .00093 - e \cos \phi + \frac{1}{2} e^2 (1 - \cos 2\phi)$$

$$\frac{15}{8} em \cos (2\xi - \phi) \tag{15}$$

$$\bar{\upsilon}(t) = \left[\dot{\Omega} \sin i \sin \eta_o + \dot{i} \cos \eta_o\right] \bar{i}_x$$

$$+ \left[\dot{\Omega} \sin i \cos \eta_o - \dot{i} \sin \eta_o\right] \bar{i}_y$$

$$+ \left[.0202 \cos 2\xi + 2e \cos \phi + \frac{15}{4} em \cos (2\xi - \phi)\right.$$

$$\left. + \frac{5}{2} e^2 \cos 2\phi\right] \bar{i}_z$$

$$= \upsilon_x \bar{i}_x + \upsilon_y \bar{i}_y + \upsilon_z \bar{i}_z \tag{16}$$

For additional details regarding the above expressions, and for an explanation of the various angular variables used, the reader is referred to Appendix B. The coordinates of the Sun in the X_e, Y_e, Z_e frame, presented in Eqs. (17) are also developed in this appendix.

The Sun's position coordinates in the rotating frame are

$$\begin{aligned} x_s &\cong r_{13} \cos \xi \\ y_s &= - r_{13} \sin \xi \\ z_s &= r_{13} \sin i \sin (\Omega - v') \end{aligned} \tag{17}$$

We now stipulate that the following quantities will be treated as being of the first order of smallness:

$$m, e, x, y, z, P_x, P_y, P_z, \sqrt{\rho(t)}, \sqrt{\upsilon(t)} \tag{18}$$

The momenta P_x, P_y, P_z conjugate to x,y,z are introduced through the relations

$$p_x = \frac{\partial L}{\partial \dot{x}} = P_x - \frac{\sqrt{3}}{2}$$

$$p_y = \frac{\partial L}{\partial \dot{y}} = P_y + \frac{1}{2} \tag{19}$$

$$p_z = \frac{\partial L}{\partial \dot{z}} = P_z$$

The terms linear in e($\cong$ .055) in $\rho(t)$ and $\upsilon(t)$ are obviously only of o(m), and will have to be treated in a different fashion if we are to retain the definition of Eq. *18). This problem will arise when we include the eccentricity in the canonical transformations to slow variables.

The use of a Taylor series to expand L and H around L_4 in terms of x,y,z, P_x, ... etc ... raises the question of how many terms of the series expansion have to be retained before we truncate it, i.e., what order of nonlinear terms must be retained so as to take into account all the dominant perturbative effects. This question is readily answered by noting that the highest internal resonance is that resulting from the near equality $\omega_1 \cong 3\omega_2$ which indicates that nonlinear terms up to and including the fourth order must be retained in the Taylor expansions of L and H.

When all the steps have been carried out and all the terms collected, as shown in Appendix C, one obtains for the Hamiltonian H, defined as usual by means of

$$H = p_r^T \dot{r} - L \tag{20}$$

where

$p_r^T = \left[p_x, p_y, p_z\right]$ = (1 x 3) row matrix of momenta elements

and

$r = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ = (3 x 1) column matrix of position elements the expression of Eq. (21)

$$
\begin{aligned}
H = H^{(o)} + H' = &\Bigg\{ \frac{1}{2}\left(P_x^2 + P_y^2 + P_z^2\right) + \left(yP_x - xP_y\right) + \frac{1}{8}\left(x^2 - 5y^2 + 4z^2\right) \\
&- \frac{3\sqrt{3}}{4}(1 - 2\mu)xy - \frac{1}{2}\dot{\rho}\left(P_x + \sqrt{3}\,P_y\right) + \frac{1}{2}\left(\rho + \upsilon_z\right)\left(\sqrt{3}\,P_x - P_y\right) \\
&+ \frac{1}{2}\left(\upsilon_y - \sqrt{3}\,\upsilon_x\right)P_z - \left(\rho + \frac{1}{2}\upsilon_z\right)\left(x + \sqrt{3}y\right) \\
&- m^2\left[\frac{3}{2r_{13}^2}\left(x_s + \sqrt{3}y_s\right)\left(x_s x + y_s y + z_s z\right) - \frac{1}{2}\left(x + \sqrt{3}y\right)\right] \\
&+ \frac{1}{2}\left(\sqrt{3}\,\upsilon_y + \upsilon_x\right)z\Bigg\}^{(o)} \\
&+ \left\{\frac{3\sqrt{3}}{16}\left(x^2y + y^3\right) + \frac{1 - 2\mu}{16}\left(33xy^2 - 7x^3 - 12xz^2\right) - \frac{3\sqrt{3}}{4}yz^2\right\}_3 \\
&+ \Bigg\{\frac{5\sqrt{3}\,(1 - 2\mu)}{32}\left(5x^3y - 9xy^3 + 12xyz^2\right) + \frac{37}{128}x^4 + \frac{3}{16}x^2z^2 + \frac{33}{16}y^2z^2 \\
&- \frac{123}{64}x^2y^2 - \frac{3}{128}y^4 - \frac{3}{8}z^4\Bigg\}_4 + \Bigg\{- \frac{3}{8}\rho\left(x^2 - 5y^2 + 4z^2 - 6\sqrt{3}\,(1 - 2\mu)xy\right) \\
&+ \upsilon_z\left(yP_x - xP_y\right) + \upsilon_y\left(xP_z - zP_x\right) + \upsilon_x\left(zP_y - yP_z\right) \\
&- m^2\left[\frac{3}{2r_{13}^2}\left(x_s x + y_s y\right)^2 - \frac{1}{2}\left(x^2 + y^2 + z^2\right)\right]\Bigg\}_s + O(m^5, m^6) \text{ etc.}
\end{aligned}
\tag{21}
$$

In the above expression we have split H' into a cubic part H_3, a quartic part H_4 and a solar part H_s, which in turn is composed of indirect solar effects (via ρ and υ) and a direct solar effect (via the m^2 term).

We shall concern ourselves in Section VII only with the motion resulting from the bracket $\{\ \}^{(o)}$ which represents the linear and quadratic part $H^{(o)}$ of H. These terms give rise to a system of forced linear differential equations which will be discussed below.

The analysis of the effect of the terms in H' on the motion of the particle will be started in Section VIII.

VII. THE LINEAR DIFFERENTIAL EQUATIONS AND THE TRANSFORMATION TO NORMAL CANONICAL COORDINATES

Hamilton's equations can be written down in a very compact form by using the matrix notation. We define the (3×1) column matrix for r and P_r in a manner similar to those introduced for r and P_r in connection with Eq. (20), and introduce the additional (1×3) row matrix of partial derivatives of $H^{(o)}$

$$
\begin{aligned}
H_r^{(o)} &= \left[\frac{\partial H^{(o)}}{\partial x}, \frac{\partial H^{(o)}}{\partial y}, \frac{\partial H^{(o)}}{\partial z}\right] \\
H_{P_r}^{(o)} &= \left[\frac{\partial H^{(o)}}{\partial P_x}, \frac{\partial H^{(o)}}{\partial P_y}, \frac{\partial H^{(o)}}{\partial P_z}\right]
\end{aligned} \tag{22}
$$

The equations of motion then can be written in the form

$$
\begin{bmatrix} \dot{r} \\ \\ \dot{P}_r \end{bmatrix} = \Phi_o \begin{bmatrix} H_{rT}^{(o)} \\ \\ H_{P_r T}^{(o)} \end{bmatrix} \qquad (6 \times 1) \text{ column matrix} \tag{23}
$$

where $H_{rT}^{(o)}$ and $H_{P_r T}^{(o)}$ are the transpose of $H_r^{(o)}$ and $H_{P_r}^{(o)}$ respectively, Φ_o is the (6×6) matrix

$$
\Phi_o = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \tag{24}
$$

I is the (3×3) identity matrix, and 0 is the (3×3) null matrix.

In component form, Eq. (23) becomes

$$
\dot{x} = H_{P_x}^{(o)} = P_x + y + \left\{-\frac{1}{2}\rho + \frac{\sqrt{3}}{2}\left(\rho + \upsilon_z\right)\right\}
$$

$$
\dot{y} = H_{P_y}^{(o)} = P_y - x + \left\{-\frac{\sqrt{3}}{2}\dot{\rho} - \frac{1}{2}\left(\rho + \upsilon_z\right)\right\}
$$

(cont. on next page)

$$\dot{z} = H^{(o)}_{P_z} = P_z + \left\{\frac{1}{2}\left(\nu_y - \sqrt{3}\,\nu_x\right)\right\}_s$$

$$\dot{P}_x = -H^{(o)}_x = P_y - \frac{1}{4}x + \frac{3\sqrt{3}}{4}(1 - 2\mu)y + \left\{\left(\rho + \frac{1}{2}\nu_z\right) + m^2\left[\frac{3}{2r_{13}^2}\left(x_s^2 + \sqrt{3}\,x_s y_s\right) - \frac{1}{2}\right]\right\}_s$$

$$\dot{P}_y = -H^{(o)}_y = -P_x + \frac{5}{4}y + \frac{3\sqrt{3}}{4}(1 - 2\mu)x + \left\{\sqrt{3}\left(\rho + \frac{1}{2}\nu_z\right) + m^2\left[\frac{3}{2r_{13}^2}\left(x_s y_s + \sqrt{3}\,y_s^2\right) - \frac{\sqrt{3}}{2}\right]\right\}_s$$

$$\dot{P}_z = -H^{(o)}_z = -z + \left\{-\frac{1}{2}\left(\sqrt{3}\,\nu_y + \nu_x\right) + m^2\left[\frac{3}{2r_{13}^2}\left(x_s z_s + \sqrt{3}\,y_s z_s\right)\right]\right\}_s \tag{25}$$

The terms in $\{\ \}_s$ contain the direct and indirect solar contributions.

The homogeneous part of Eq. (25) is obtained by setting $\rho = \nu = m = o$. The characteristic equation resulting from a trial solution $e^{i\omega t}$ is

$$\left(1 - \omega^2\right)\left[\omega^4 - \omega^2 + \frac{27}{16} - \eta^2\right] = 0 \tag{26}$$

where

$$\eta = \frac{3\sqrt{3}}{4}(1 - 2\mu) = 1.26753$$

The solutions to Eq. (26) are the eigenvalues

$$\begin{aligned} \omega_1 &= \pm .95459 \\ \omega_2 &= \pm .29791 \\ \omega_3 &= \pm 1.0 \quad \text{(corresponds to an uncoupled z motion)} \end{aligned} \tag{27}$$

The above ω's are the natural frequencies which were used in the discussion of the detunings in Section III.

Let the solutions of the homogeneous set of equations be denoted by $\bar{x}, \bar{y}, \bar{z} \cdots$ and suitable particular integrals by $\tilde{x}, \tilde{y}, \tilde{z} \cdots$ Thus the complete solutions are

$$
\begin{aligned}
x &= \bar{x} + \tilde{x} \\
y &= \bar{y} + \tilde{y} \\
\vdots & \quad \vdots \quad \vdots
\end{aligned}
\tag{28}
$$

For later use the 6 constants of integration which appear in the solutions (28) are best introduced by transforming first to a normal canonical set of coordinates Q and momenta P

$$
Q^T = \left[Q_1,\ Q_2,\ Q_3\right] \qquad P^T = \left[P_1,\ P_2,\ P_3\right] \tag{29}
$$

which also satisfy Hamilton's equations of motion and represent uncoupled motions in the form of independent simple harmonic oscillations having as frequencies the three eigenvalues ω_i.

The linear equations of transformation can be written in the form

$$
\begin{pmatrix} \bar{x} \\ \bar{y} \\ \bar{z} \\ \bar{P}_x \\ \bar{P}_y \\ \bar{P}_z \end{pmatrix} = J \begin{pmatrix} Q_1 \\ Q_2 \\ Q_3 \\ P_1 \\ P_2 \\ P_3 \end{pmatrix} \tag{30}
$$

where J is a (6 $\times$ 6) matrix whose columns consist of the eigenvectors corresponding to the eigenvalues $\pm\,\omega_1$, and which are normalized so as to satisfy Eq. (31) which is the necessary condition for a canonical transformation.

$$J\ \Phi_o\ J^T = \Phi_o \tag{31}$$

The matrix presented in Eq. (32) satisfies this conditions and thus provides the proper coordinate transformation.

$$J = \begin{bmatrix} 0 & 0 & 0 & \frac{K_1}{\omega_1}\left(\omega_1^2 + \frac{9}{4}\right) & -\frac{K_2}{\omega_2}\left(\omega_2^2 + \frac{9}{4}\right) & 0 \\ -2K_1\omega_1 & -2K_2\omega_2 & 0 & -\frac{K_1}{\omega_1}\eta & \frac{K_2}{\omega_2}\eta & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -K_1\omega_1\left(\omega_1^2 + \frac{1}{4}\right) & -K_2\omega_2\left(\omega_2^2 + \frac{1}{4}\right) & 0 & \frac{K_1}{\omega_1}\eta & -\frac{K_2}{\omega_2}\eta & 0 \\ K_1\omega_1\eta & K_2\omega_2\eta & 0 & \frac{K_1}{\omega_1}\left(\frac{9}{4} - \omega_1^2\right) & -\frac{K_2}{\omega_2}\left(\frac{9}{4} - \omega_2^2\right) & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \end{bmatrix} \tag{32}$$

where

$$K_i = \left\{\left|\frac{11}{2}\omega_i^2 + 2\eta^2 - \frac{45}{8}\right|\right\}^{-1/2} \qquad i = 1,2$$

$$K_1 = .62016$$

$$K_2 = .72101$$

The numerical values of the elements in J are

$$J = \begin{bmatrix} 0 & 0 & 0 & 2.05374 & -5.66028 & 0 \\ -1.24032 & -1.44202 & 0 & -.823463 & 3.06768 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -.687459 & -.0727629 & 0 & .823463 & -3.06768 & 0 \\ .750378 & .272262 & 0 & .869732 & -5.23066 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \end{bmatrix} \tag{33}$$

In terms of Q and P the Hamiltonian $H^{(o)}$ (for the case $\rho = \upsilon = m = o$) becomes

$$H^{(o)} = \frac{1}{2}\left(P_1^2 + \omega_1^2 Q_1^2\right) - \frac{1}{2}\left(P_2^2 + \omega_2^2 Q_2^2\right) + \frac{1}{2}\left(P_3^2 + \omega_3^2 Q_3^2\right) \qquad (34)$$

The solutions for the three harmonic oscillators which make up the expression for $H^{(o)}$ in Eq. (34) can be given in the form

$$\begin{aligned} Q_1 &= \frac{\sqrt{2\alpha_1}}{\omega_1} \sin \omega_1 \beta_1^{\neq} \\ Q_2 &= \frac{\sqrt{2\alpha_2}}{\omega_2} \sin \omega_2 \beta_2^{\neq} \\ Q_3 &= \frac{\sqrt{2\alpha_3}}{\omega_3} \sin \omega_3 \beta_3^{\neq} \\ P_1 &= \sqrt{2\alpha_1} \cos \omega_1 \beta_1^{\neq} \\ P_2 &= -\sqrt{2\alpha_2} \cos \omega_2 \beta_2^{\neq} \\ P_3 &= \sqrt{2\alpha_3} \cos \omega_3 \beta_3^{\neq} \end{aligned} \qquad (35)$$

where $\beta_1^{\neq} = t + \beta_1$, $\beta_2^{\neq} = t - \beta_2$, $\beta_3^{\neq} = t + \beta_3$, and α_i, β_i are the 6 required constants of integration.

Substitution of Eq. (35) and the J matrix (33) into Eq. (30) gives the homogeneous solutions for the coordinates

$$\begin{aligned} \bar{x} &= 2.902 \sqrt{\alpha_1} \cos \omega_1 \beta_1^{\neq} + 8.003 \sqrt{\alpha_2} \cos \omega_2 \beta_2^{\neq} \\ \bar{y} &= 2.103 \sqrt{\alpha_1} \cos \left(\omega_1 \beta_1^{\neq} + 123.57^{o}\right) \\ &\quad + 4.793 \sqrt{\alpha_2} \cos \left(\omega_2 \beta_2^{\neq} + 154.82^{o}\right) \\ \bar{z} &= \sqrt{2\alpha_3} \cos \beta_3^{\neq} \end{aligned} \qquad (36)$$

The particle trajectories in the xy plane corresponding to each of the two coplanar normal modes are ellipses with major axes at right angles to the vector $\bar{r}_{1L}$ and thickness ratios (minor axis/major axis) 1:2 for ω_1 and 1:5 for ω_2 as shown in Fig. 5. Motion proceeds in a clockwise direction.

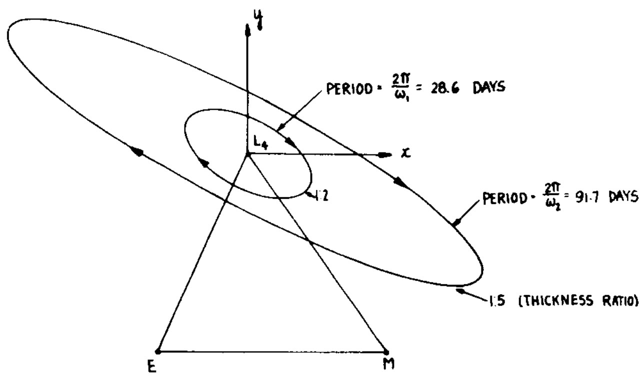


Fig. 5: Trajectories of normal modes.

The complete unperturbed xy motion consists of a weighted superposition of these two normal modes, and is in general not periodic.

The particular solutions $\tilde{x},\tilde{y}$, corresponding to the forcing functions contained in the { } brackets of Eq. (25) are most readily obtained from the coplanar equations

$$\begin{aligned} \ddot{x} - 2\dot{y} - \frac{3}{4}x - \eta\, y &= \dot{f}_x - f_y + f_{P_x} \\ \ddot{y} + 2\dot{x} - \eta\, x - \frac{9}{4}y &= \dot{f}_y + f_x + f_{P_y} \end{aligned} \tag{37}$$

where f_x, f_y, f_{p_x}, f_{p_y} denote the direct and indirect solar forcing functions of the subscript variables given in equations (25). For our purposes it is sufficient to obtain the particular solutions to $o(m^2)$. After introducing Eqs. (16) into (37) we obtain the solutions

$$\begin{aligned} \tilde{x} &= \tilde{x}_o + \tilde{x}_e + \tilde{x}_{em} + \tilde{x}_{e^2} + \tilde{x}_{m^2} + \tilde{x}_c \\ \tilde{y} &= \tilde{y}_o + \tilde{y}_e + \tilde{y}_{em} + \tilde{y}_{e^2} + \tilde{y}_{m^2} + \tilde{y}_c \end{aligned} \tag{38}$$

where

$$\left.\begin{aligned} \tilde{x}_o &= .01016 \cos (2\xi - 67.2^o) \\ \tilde{y}_o &= .00867 \cos (2\xi + 38.3^o) \end{aligned}\right\} \quad \text{Resulting from the indirect solar terms}$$

$$\begin{aligned} \tilde{x}_e &= .31\, e \cos (\phi - 72.2^o) \\ \tilde{y}_e &= .227\, e \cos (\phi + 50.2^o) \\ \tilde{x}_{em} &= 11.1\, em \cos (2\xi - \phi - 75.2^o) \\ \tilde{y}_{em} &= 7.86\, em \cos (2\xi - \phi + 51.76^o) \\ \tilde{x}_{e^2} &= 1.274\, e^2 \cos (2\phi + 30.8^o) \\ \tilde{x}_{e^2} &= 1.062\, e^2 \cos (2\phi - 66.0^o) \end{aligned}$$

$$\left.\begin{aligned} \tilde{x}_{m^2} &= 1.697\, m^2 \cos (2\xi - 127.7^o) \\ \tilde{y}_{m^2} &= 1.43\, m^2 \cos (2\xi - 20.83^o) \end{aligned}\right\} \quad \text{Resulting from the direct solar terms}$$

$$\left.\begin{aligned} \tilde{x}_c &= .50\, e^2 \\ \tilde{y}_c &= .2895\, e^2 \end{aligned}\right\} \quad \text{Constant displacement}$$

No particular solution for $\tilde{z}$ is retained since it is of $o(m^3)$ or higher, and would lead to terms of $o(m^5)$ when substituted into H'. On this point we shall have something more to say in Section XIII.

The corresponding solutions for $\tilde{P}_r$ are readily obtained from the relations

$$\begin{aligned} P_x &= \dot{x} - y \\ P_y &= \dot{y} + x \\ &\text{etc.} \end{aligned} \tag{39}$$

It is interesting to note that if we substitute Eq. (35) into Eq. (34) we obtain the simple expression

$$H^{(o)} = \alpha_1 - \alpha_2 + \alpha_3 \tag{40}$$

The particular manner of introducing the polar set of integration constants α_i, β_i into Eq. (35) follows from the canonical relationship which they bear the Hamiltonian $H^{(o)}$. The quantities $\beta_1^{\neq}$, $-\beta_2^{\neq}$, $\beta_3^{\neq}$ and α_1, α_2, α_3 form, respectively, a canonical set of coordinates and conjugate momenta with respect to $H^{(o)}$ of Eq. (40).

Thus

$$\begin{aligned} \dot{\beta}_1^{\neq} &= 1 = H^{(o)}_{\alpha_1} & \dot{\alpha}_1 &= -H^{(o)}_{\beta_1^{\neq}} = 0 \quad \text{or} \quad \alpha_1 = \text{const} \\ \dot{\beta}_2^{\neq} &= -1 = H^{(o)}_{\alpha_2} & \dot{\alpha}_2 &= +H^{(o)}_{\beta_2^{\neq}} = 0 \quad \text{or} \quad \alpha_2 = \text{const} \\ \dot{\beta}_3^{\neq} &= 1 = H^{(o)}_{\alpha_3} & \dot{\alpha}_2 &= -H^{(o)}_{\beta_3^{\neq}} = 0 \quad \text{or} \quad \alpha_3 = \text{const} \end{aligned} \tag{41}$$

The above results are in agreement with our stipulation that α_i and β_i be constants.

Furthermore, the quantities α_i and β_i themselves form a canonical set with respect to an unperturbed Hamiltonian $H = 0$.

The above canonical properties will be made use of when we analyze the perturbative effect of H'.

The form of $H^{(o)}$ in Eq. (40) makes it very easy to verify the point made earlier in B(2) of Section III, regarding the sign indeterminacy of H which is seen to depend, for small α_3, on the difference

$\alpha_1 - \alpha_2$. Although in the present case α_1 and α_2 individually are constants, it turns out that for the case $\alpha_3 \equiv 0$ the combination $\alpha_1(t) - \alpha_2(t)$ remains a constant of the motion also when the perturbative effects of higher order internal nonlinearities are included (but external solar perturbations and lunar eccentricity are still neglected). Thus α_1 and α_2 may grow individually as long as their difference remains fixed, which indicates the possibility of an internally generated instability near L_4 also for the classical restricted 3-body problem (for which we use the exact expression for H).

That H is a constant of the motion in the latter case (where $H \neq H(t)$) as stated in A(4) of Section III, is readily verified since

$$\frac{dH}{dt} = \frac{\partial H}{\partial r}\,\dot{r} + \frac{\partial H}{\partial p}\,\dot{p} = \text{using Eq. (23)} = -\,\dot{p}^T\dot{r} + \dot{r}^T\dot{p} = 0 \qquad (42)$$

The only existing integral of the motion, the Jacobi constant C_J (see pp. 281 of Ref. 9 where it is denoted by unsubscripted C), is equal to the negative of H

$$H = -\,C_J \qquad (43)$$

VIII. MODIFICATION OF THE LINEAR SOLUTION DUE TO H'

The inclusion of the terms in the Hamiltonian H', neglected until now in the previous solution, can be handled by a method equivalent to the customary variations of constants technique by requiring the original constants of integration α and β introduced in Eq. (35) to become functions of time, which then satisfy Hamilton's equations with Hamiltonian H'.

Inasmuch as we are not concerned in the present investigation with an exact or detailed determination of the particle's trajectory, but rather in the overall broad features of the motion, we shall desire to obtain only the slowly varying components of α and β which will arise from the secular terms in H', and those terms containing low combination frequencies which arise from the near resonances.

This can be accomplished by means of a suitable canonical transformation of coordinates from the polar canonical set α,β associated with $H = 0$ to a new slowly varying canonical set α',β' associated with a new slowly varying Hamiltonian K'. K' will contain only the lowest frequency terms which arise in H' as a result of the above transformation, all other faster terms having been suitably eliminated. The question as to which frequencies should be retained, and the cut-off point beyond which the periodic terms are dropped cannot be readily answered in general terms, but would depend on the particular problem considered, and also on the density of spacing of the resonance peaks in the lower end of the frequency spectrum. This point will be touched upon again later in connection with the specific form of the expression for K'.

Returning once more to the coordinate transformation mentioned earlier, it is reasonable to assume that for relatively small displacements x,y,z of the particle, the effect of H' would be in the nature of a perturbation of the linearized solution found earlier. With this assumption in mind we may now consider a stationary contact transformation

$$\begin{aligned} \alpha_i' &= \alpha_i + \delta\alpha_i \\ \beta_i' &= \beta_i + \delta\beta_i \end{aligned} \tag{44}$$

that may be introduced with the aid of a generating function $G(\beta,\alpha')$

$$G(\beta,\alpha') = \beta\alpha' + S(\beta,\alpha') \tag{45}$$

which satisfies the relations[10]

$$\begin{aligned} \beta' &= \frac{\partial G}{\partial \alpha'} = \beta + S_{\alpha'} \\ \alpha &= \frac{\partial G}{\partial \beta} = \alpha' + S_{\beta} \end{aligned} \tag{46}$$

The first term $\beta\alpha'$ in G generates the identity transformation, while the function $S(\beta,\alpha') = S_1 + S_2$ denotes an additional suitably selected generating function which is introduced for the specific purpose of eliminating all the short period terms which occur in H': S_1 is selected to eliminate the terms of $o(m^3)$ and S_2 those of $o(m^4)$.

Since S does not depend explicitly on time t we can write

$$K'(\beta',\alpha',t) = H(\beta',\alpha',t) + H'(\beta',\alpha',t) \tag{47}$$

where H above is evaluated in terms of the new coordinates β' and new momenta α'.

When all the required steps of the transformation are carried out, as indicated in Appendix D, one arrives at the following relation for K'

$$K' = \tilde{\bar{H}}_3 + \tilde{\bar{H}}_4 + \bar{H}_4 - \frac{1}{2}\overline{\left[H_3,\ S_1\right]} \tag{48}$$

$\tilde{\bar{H}}_3$ and $\tilde{\bar{H}}_4$ are the long period terms resulting from the Taylor series expansions

$$\tilde{\bar{H}}_i = \sum_{n=1}^{\infty} \frac{1}{n!} \left[\tilde{x}\frac{\partial}{\partial x} + \tilde{y}\frac{\partial}{\partial y} + \tilde{z}\frac{\partial}{\partial z}\right]^n H_i \tag{49}$$

evaluated at $\bar{x},\bar{y},\bar{z}$.

$\overline{[H_3,S_1]}$ denotes the long period part of the Poisson bracket of H_3 with S_1. $\bar{H}_4$ results from the substitution of the homogeneous solutions $\bar{x},\bar{y},\bar{z}$ into H_4, and consists of an internal part $\bar{H}_{4int}$ and an external part $\bar{H}_{4ext}$ which contains both the direct and indirect solar effects.

The algebraic work needed to express K' in terms of α',β' and t is rather formidable, and is one of the major stumbling blocks in what would otherwise be a relatively straightforward solution. A few representative steps of the required manipulations are briefly demonstrated in Appendix E. If all the manipulations have been successfully carried out, one does eventually come up with an expression for K' which has the general form shown in Eq. (50).

$$\begin{aligned}K' = {} & \alpha_1'\left[b_1 + b_2 C_{2\Delta\xi_1+\lambda_1}\right] + \alpha_2' b_3 + \alpha_3' b_4 + \alpha_1'^{3/2} b_5 C_{\Delta\phi_1+\lambda_2} \\ & + \alpha_2'^{3/2}\left[b_6 C_{\Delta\phi_2+\lambda_3} + b_7 C_{\sigma-\Delta\phi_1+\lambda_4}\right] + \alpha_1'^{1/2}\alpha_2\left[b_8 C_{\Delta\phi_1+\lambda_5}\right] \\ & + \alpha_1'^{1/2}\alpha_3'\left[b_9 C_{\Delta\phi_1+\lambda_6} + b_{10} C_{\Delta_{13}+\Delta\phi_3+\lambda_7}\right] + \alpha_1'\alpha_2' b_{11} + \alpha_1'^2 b_{12} \\ & + \alpha_2'^2 b_{13} + \alpha_2'^{1/2} b_{14} C_{\sigma+\lambda_8} + \alpha_1'\alpha_2'^{1/2}\left[b_{15} C_{\sigma-\Delta\phi_1+\lambda_9} + b_{16} C_{\Delta\phi_2+\lambda_{10}}\right] \\ & + \alpha_3'^2 b_{17} + \alpha_1'\alpha_3'\left[b_{18} + b_{19} C_{2\Delta_{13}+\lambda_{11}}\right] + \alpha_2'\alpha_3' b_{20} \\ & + \alpha_2'^{1/2}\alpha_3'\left[b_{21} C_{\Delta\phi_2+\lambda_{12}} + b_{22} C_{\Delta_{13-\sigma+\Delta\phi_3+\lambda_{13}}}\right]\end{aligned} \tag{50}$$

$b_j (j = 1,\cdots 22)$ are known constants and C_x stands for cos x. The detuning frequencies retained in Eq. (50) have the following magnitudes.

$$\begin{aligned}
&2\Delta\xi_1 \rightarrow 2(1 - m - \omega_1) = -.05878 \\
&2\Delta\phi_1 \rightarrow 2(1 - \tfrac{3}{4}m^2 - \omega_1) = .08242 \\
&\Delta\phi_2 = \phi - 3\omega_2\beta_2^{\neq} \rightarrow (1 - .0042 - 3\omega_2) = .10207 \\
&\sigma = \omega_1\beta_1^{\neq} - 3\omega_2\beta_2^{\neq} \rightarrow \omega_1 - 3\omega_2 = .06086 \\
&\Delta_{13} \rightarrow \omega_1 - 1 = -.04541 \\
&\Delta_{13} + \Delta\phi_3 \rightarrow -.04541 - .0042 = -.04961 \\
&\sigma - \Delta\phi_1 \rightarrow .06086 - .04121 = .01965
\end{aligned} \tag{50a}$$

and

$$\Delta_{13} - \sigma + \Delta\phi_3 \rightarrow -.1105$$

The terms containing $\Delta\phi$ arise from the lunar eccentricity.

As can be seen from Eqs. (50a) no terms with frequencies larger than .12 have been retained in the expression for K'. Although this choice of cut-off frequency appears at first sight rather arbitrary, it can be argued here that for higher frequencies the resultant detuning would not be narrow enough to introduce the very small divisors which usually lead to divergent solutions, and that consequently their omission should not materially affect the overall features of the resultant particle motion.

The large number of frequencies which still are left in K' pose considerable difficulties in the way of a straightforward analytical treatment. To enable one to carry out nonetheless a reasonably meaningful analysis of the effects of internal resonances and of the solar perturbation, it was found necessary to reduce the number of admissible resonance peaks still further. This was accomplished by disregarding for the present time from further consideration all the terms which arise from the lunar orbital eccentricity. While this step does tend to restrict the present analysis to encompass only circular lunar orbits, it manages to reduce the number of detuning frequencies left down to 3. For this number of resonances an analysis can be carried out.

Eccentricity terms could perhaps be reintroduced at a later time, possibly by means of an additional perturbation of the variational equations which result from the present circular orbit analysis. A possible shortcoming with such a scheme might be that it would probably lead to a set of parametrically excited linear differential equations which would not be readily solvable.

Another somewhat different approach might be attempted, if we recall that the elliptic 3-body problem (no solar perturbation present) admits as a solution an elliptic particle orbit around L_4. This ellipse is identical to the ellipse along which the moon appears to move relative to an observer moving with constant circular velocity along the moon's mean circular reference orbit, but rotated 60° with respect to it. Stated another way, the particle's motion is synchronized with that of the moon, but takes place 60° ahead of it. Variational equations for these orbital elements due to the solar perturbation could then be set up and hopefully solved.

The above are just two of the many other different approaches which might have to be explored in greater detail before the more general question of stability of motion could be satisfactorily resolved.

In the present dissertation however, we shall hereafter confine our attention only to the case of zero lunar eccentricity.

IX. THE LONG PERIOD HAMILTONIAN FOR e = o AND THE ELIMINATION OF TIME t

For the case of e = o the expression for K' shown in Eq. (50) is reduced to the simpler form given in Eq. (51) below. The numerical values of the coefficients b, and the phase shifts λ, are determined after one performs all the tedious algebraic manipulations similar to those briefly demonstrated in Appendix E. There results

$$
\begin{aligned}
K' = \Big\{ & .1266\alpha_1'^2 - 6.000\alpha_1'\alpha_2' + 3.829\alpha_2'^2 \\
& - 29.04\alpha_1'^{1/2}\alpha_2'^{3/2} \cos\left[.06086t + \omega_1\beta_1' + 3\omega_2\beta_2' + 14.2^\circ\right] \\
& + \alpha'\alpha_3'\left[.09316 + .08608 \cos 2\Delta_{13} - .03934 \sin 2\Delta_{13}\right] \\
& + .7554\alpha_2'\alpha_3' - .002231\,\alpha_3'^2\Big\}_{int} - \Big\{.005394\alpha_1' + .008208\alpha_2' \\
& + .02685\alpha_1' \cos\left[.05878t + 2\omega_1\beta_1' + 29.4^\circ + 2\epsilon' - 2\epsilon\right] + .004193\alpha_3'\Big\}_{ext}
\end{aligned}
\tag{51}
$$

where

$$\Delta_{13} = \omega_1(t + \beta_1') - (t + \beta_3') = -.04541t + \omega_1\beta_1 - \beta_3'$$

The first bracket contains all the internal terms, while the second bracket includes all the external (solar) terms. The long period contributions to the coplanar (α_1',α_2') terms resulting from the periodic parts of the indirect $\rho(t)$ and $\upsilon(t)$ terms in H' were found to cancel exactly the indirect periodic terms generated by the linear forced response $\tilde{x}_o$ and $\tilde{y}_o$ of Eq. (38). The external terms displayed in Eq. (51), which are left after the above cancellations, stem from the contribution of the indirect constant component -.00093 in ρ, from the direct (m^2) terms in H, and from the forced responses $\tilde{x}_{m^2}$ and $\tilde{y}_{m^2}$ of the linear system.

Equation (51) shows that the dependence of K' on time t comes about through the presence of three distinct slowly varying trigonometric terms with frequencies .06086, .09082 and .05878, all of which are of o(m). Since the same trigonometric functions also depend on various combinations of the three angular variables $\beta_i'(i = 1,2,3)$, the possibility suggests itself to eliminate the explicit presence of t by means of a suitable redefinition of the β_i' so as to absorb the time dependent terms. Such a transformation would result in a new Hamiltonian K^* which would not depend explicitly on t.

This absorption of the time terms is accomplished by means of a coordinate transformation to a new canonical set of variables α^* and β^* as indicated below.

We define β_1^* via

$$2\beta_1^* = .05878t + 2\omega_1\beta_1' + 29.4^\circ - 2\epsilon + 2\epsilon'$$

or

$$\beta_1^* = .02939t + \omega_1\beta_1' + 14.7^\circ - \epsilon + \epsilon' \tag{52}$$

The conjugate momentum α_1^* is obtained by the introduction of a generating function J_1 defined as

$$J_1 = \alpha_1^*\left[.02939t + \omega_1\beta_1' + 14.7^\circ - \epsilon + \epsilon'\right] \tag{53}$$

so that

$$\alpha_1' = \frac{\partial J_1}{\partial \beta_1'} = \omega_1\alpha_1^* \qquad \text{or} \qquad \alpha_1^* = \frac{\alpha_1'}{\omega_1} \tag{54}$$

For the definition of β_2^* we use the trigonometric argument

$$.06086t + \omega_1\beta_1' + 3\omega_2\beta_2' + 14.2^\circ$$

and substitute for β_1' from Eq. (52). This leads to the expression

$$\left[.03146t + 3\omega_2\beta_2' + \epsilon - \epsilon' \quad .5^{\circ}\right] + \beta_1^*$$

which suggests that β_2^* be taken as

$$\beta_2^* = .03146t + 3\omega_2\beta_2' + \epsilon - \epsilon' - .5^{\circ} \tag{55}$$

Use of a second generating function

$$J_2 = \alpha_2^*\left[.03146t + 3\omega_2\beta_2' + \epsilon - \epsilon' - .5^{\circ}\right] \tag{56}$$

gives for the conjugate momentum α_2^*

$$\alpha_2^* = \frac{\alpha_2'}{3\omega_2} \tag{57}$$

The expressions for β_3^* and α_3^* can be obtained in a similar fashion with the aid of Δ_{13}. Combining first the cosine and sine terms

$$.08608 \cos 2\Delta_{13} - .03934 \sin 2\Delta_{13} = .09464 \cos\left[2\Delta_{13} + 24.56^{\circ}\right]$$

we find that

$$\beta_3^* = .074801t + \omega_3\beta_3' - \epsilon + \epsilon' + 2.42^{\circ} \tag{58}$$

and after introducing a generating function J_3 we obtain

$$\alpha_3^* = \frac{\alpha_3'}{\omega_3} = \alpha_3' \tag{59}$$

Letting $J = J_1 + J_2 + J_3$ and noting that $J = J(\beta', \alpha^*, t)$ we determine the transformed time independent Hamiltonia K^* from the relation

$$K^* = K'(\beta^*, \alpha^*) + \frac{\partial J(\beta', \alpha^*, t)}{\partial t} \tag{60}$$

Substitution for α', β' in terms of α^*, β^* in Eq. (51) and use of Eq. (60) results in the desired expression for K^*:

$$\begin{aligned} K^* = \Big\{ & .1154\alpha_1^{*2} - 5.1\alpha_1^*\alpha_2^* + 3.059\alpha_2^{*2} - 23.97\alpha_1^{*1/2}\alpha_2^{*3/2}\, C_{\beta_1^*+\beta_2^*} \\ & + .09035\alpha_1^*\alpha_3^* C_{2(\beta_1^*-\beta_3^*)} + .08893\alpha_1^*\alpha_3^* + .6751\alpha_2^*\alpha_3^* \\ & - .002231\alpha_3^{*2} + .02939\alpha_1^* + .03146\alpha_2^* + .074801\alpha_3^* \Big\}_{\text{int}} \\ & + \Big\{ .004193\alpha_3^* - .007336\alpha_2^* - \alpha_1^*\Big[.005149 + .02563 C_{2\beta_1^*} \Big] \Big\}_{\text{ext}} \end{aligned} \tag{61}$$

where the notation $C_x \equiv \cos x$ has again been used for convenience.

X. ANALYSIS OF THE INTERNAL COPLANAR MOTION

1. SIMPLIFICATION OF THE HAMILTONIAN

The analysis of the motion governed by the Hamiltonian K^* of Eq. (61) is made easier, and a greater amount of physical insight is gained, if we treat at first separately the internal terms contained in the first bracket. The modifications required by the presence of the second, external, bracket are then taken up later.

Let us write for convenience

$$K^* = K_i^* + K_e^* \tag{62}$$

where

K_i^* = all the internal terms

K_e^* = all the external terms

and confine our attention in this and the next section to the Hamiltonian K_i^*.

It would help matters appreciably if we could eliminate also for the time being the coupling which exists between the out-of-plane and coplanar terms.

This elimination can be accomplished by a suitable choice of initial conditions which result in $\alpha_3^* \approx 0$, provided we have reason to believe that a physical motion in which α_3^* does not depart much from its initial small value can in fact exist.

The resultant coplanar type of motion can be maintained as long as the nonlinear coupling with the out-of-plane terms does not lead to an appreciable transfer of energy from one mode of motion to the other.

In the next section, where we consider the out-of-plane motion, this situation will be shown to hold true.

On the basis of the foregoing we shall neglect here all the α_3^* terms in K^*, which leaves us with the 2-dimensional Hamiltonian K_{12}^* given by

$$K_{i2}^* = .1154\alpha_1^{*2} - 5.1\alpha_1^*\alpha_2^* + 3.059\alpha_2^{*2} - 23.97\alpha_1^{*1/2}\alpha_2^{*3/2}C_{\beta_1^*+\beta_2^*}=$$

$$+ .02939\alpha_1^* + .03146\alpha_2^* \tag{63}$$

Since t is not explicitly present in K_{i2}^*, the latter can also be treated as a constant of the motion.

2. INVARIANCE OF THE DIFFERENCE $\alpha_1^* - \alpha_2^*$ AND BOUNDED MOTIONS

The presence of β_1^* and β_2^* in K_{i2}^* occurs only through the combination $\beta_1^* + \beta_2^*$. From this one readily sees that

$$\frac{\partial K_{i2}^*}{\partial \beta_1^*} = \frac{\partial K_{i2}^*}{\partial \beta_2^*}$$

which implies that

$$\dot{\alpha}_1^* = \dot{\alpha}_2^* \tag{64}$$

and after integration results in the additional coplanar integral of the motion

$$\alpha_1^* - \alpha_2^* = D_1 = \pm |D_1| \tag{65}$$

Unfortunately, this last integral does not provide any bounds on the magnitude of the coplanar displacements, inasmuch as α_1^* and α_2^* are not prohibited by Eq. (65) from growing individually as long as their difference remains unchanged.

On the other hand it is clear that the validity of the present fourth order theory would cease to hold long before the α's have grown to very large size, and that additional higher order terms in H would have to be included in the analysis. Equations (64) and (65) are of great use in those cases when α_1^* and α_2^* do not grow without limit.

Let us consider now the question of boundaries of α^*. From Eq. (64) we have

$$\left(\dot{\alpha}_2^*\right)^2 = \left(-\frac{\partial K_{i2}^*}{\partial \beta_2^*}\right)^2 = 23.97^2 \alpha_1 \alpha_1^* \alpha_2^{*3} S^2_{\beta_1^*+\beta_2^*}$$

$$= 23.97^2 \alpha_1^* \alpha_2^{*3} - \left[K_{i2}^* - .1154\alpha_1^{*2} + 5.1\alpha_1^* \alpha_2^* - 3.059\alpha_2^{*2} - .02939\alpha_1^* - .03146\alpha_2^*\right]^2 \tag{66}$$

We now introduce the new variable

$$\xi = \frac{\alpha_2^*}{|D_1|} \tag{67}$$

and Eq. (65) into Eq. (66), which can then be written in the form

$$\left(\frac{\xi}{23.97 D_1}\right)^2 = f^2(\xi) - \eta^2(\xi) \tag{68}$$

where

$$f = \begin{cases} \pm\left[\xi^3(\xi + 1)\right]^{1/2} & \text{for} \quad D_1 > 0 \\ \pm\left[\xi^3(\xi - 1)\right]^{1/2} & \text{for} \quad D_1 < 0 \end{cases} \tag{69}$$

and

$$\eta = \begin{cases} -\left(.2028 - \dfrac{.00254}{|D_1|}\right)\xi - .0801\,\xi^2 + \text{constant} & D_1 > 0 \\ \left(.2028 + \dfrac{.00254}{|D_1|}\right)\xi - .0801\,\xi^2 + \text{constant} & D_1 < 0 \end{cases} \tag{70}$$

The constants in Eq. (70) denote the value of $\eta(o)$ and are related to the value of the Hamiltonian K^*_{i2}.

The points at which the η curve intersects the + or - branch of the f curve

$$\eta = \pm f \tag{71}$$

correspond to points at which $\dot{\alpha}^*_2 = 0$ and, by Eq. (65), also $\dot{\alpha}^*_1 = 0$. Reality of the particle motions requires that $f^2 \geq \eta^2$.

The gradual changes of the motion of the physical particle in the xy space can be described by observing the motion of a representative mathematical point along a given curve η in a plane in which f and η are plotted as functions of ξ.

If the η curve intersects both branches of the f curve or intersects the same branch at two different points, then $\dot{\alpha}^*_1$ and $\dot{\alpha}^*_2$ will have finite values at intermediate points on η, which tend toward zero as the representative point approaches the f curve. The sense of motion of the point is reversed every time one of the branches of f is reached, so that the point continues to travel back and forth on a given η curve between its points of intersection with f. The turning or extremal values of the momenta α^* are thus fixed by the values whic ξ assumes at the points of intersection of η with $\pm$ f.

The geometry in the $f(\xi)$ and $\eta(\xi)$ plane is shown schematically in Fig. 6.

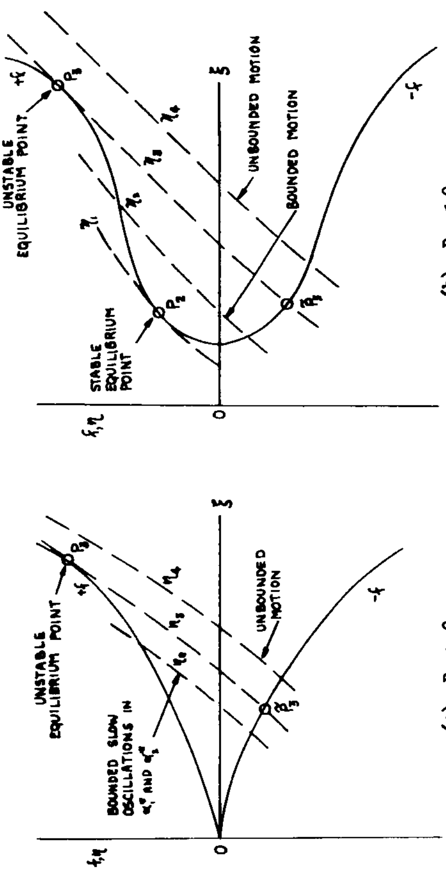
The curves η_2 in Figs. 6(a) and (b) represent bounded particle trajectories in the xy plane. The tangency points P_2, P_3 at which

$$\frac{d}{d\xi}\eta \equiv \eta' = \pm f'$$

and

$$\dot{\alpha}^*_1 = \dot{\alpha}^*_2 = 0 \tag{72}$$

are equilibrium points in the (α^*_1, α^*_2) plane, and with the aid of Eq. (66) can be shown to correspond to coplanar periodic particle



(a) $D_1 > 0$

(b) $D_1 < 0$

Fig. 6: Geometry of coplanar internal motion in α^* Space.

orbits. Equation (66) requires that $\beta_1^* + \beta_2^* = n\pi$, which can also be written in the form

$$\omega_1\beta_1' + 3\omega_2\beta_2' + .06086t - n\pi + 14.2^\circ = 0$$

Reference to Eq. (51) shows that this condition eliminates the detuning term due to coplanar coupling and indicates commensurability of the internally perturbed coplanar normal frequencies $\omega_1' = \omega_1 + \omega_1\dot{\beta}_1'$ and $\omega_2' = \omega_2 + \omega_2\dot{\beta}_2'$. The periodicity of the coplanar particle orbits follows from here.

The equilibrium is stable at point P_2 and unstable at point P_3, where small disturbances may cause a displacement to a neighboring curve such as η_4 which causes divergence of the physical motion.

Transition from stability to instability occurs at points where

$$\eta'' = \pm f'' \tag{73}$$

When $D_1 > 0$, f'' does not change sign as can be seen in Fig. 6(a), and from this follows that all the periodic particle orbits for which $\alpha_1^* > \alpha_2^*$ would be of the unstable kind. For the case $D_1 < 0$, f'' does change sign at some value $\xi > 1$ and we note accordingly the presence of one stable and one unstable equilibrium point along the +f branch in Fig. 6(b).

3. THE PERIODIC MOTIONS

When one solves the tangency Eq. (72) for the value of ξ which corresponds to every choice of D_1, one can obtain an α_1^* for every α_2^* found. In the α_1^* versus α_2^* plane this solution curve represents the so called "tangency locus" of equilibrium values of α_1^* and α_2^* which designate periodic particle orbits. This curve is presented in Fig. 7, where we have chosen as coordinates the quantities $10\sqrt{\alpha_1^*}$ and $10\sqrt{3\alpha_2^*}$ (α_1^* and $3\alpha_2^*$ are in fact the associated "action variables"). On this curve we have set the angular variables Δ_{12}^*

$$\Delta_{12}^{*} = \beta_1^{*} + \beta_2^{*} = \begin{cases} 0 \\ \text{or} \\ \pi \end{cases} \tag{74}$$

This plot is seen to consist of two distinct branches which connect at the point (1.12,0). The left hand branch consists of a segment of stable periodic orbits which is followed by a segment of unstable periodic orbits. On both segments $\Delta_{12}^{*} = 0$. The unstable branch on the right hand side of (1.12,0) requires a $\Delta_{12}^{*} = \pi$.

Two more curves passing through (1.12,0) and consisting of left hand and right hand branches are also shown in this figure. The lower (solid) curve denotes the loci of intersection points $\tilde{P}_3$ of η_3 with the second f branch. (For added clarification small inserts of the appropriate geometrical situation described by Fig. 6 are also displayed here in connection with specific segments of the curves.)

The dashed curve lying close to the $\tilde{P}_3$ locus represents the intersection of η_3 with the ξ axis. On this curve $\Delta_{12}^{*} = \pi/2$. The values of Δ_{12}^{*} which allow stable motions to exist in each one of the domains I - IV which are separated by the above curves are indicated in the figure, and also by shaded regions in the small inserts from Fig. (6).

The axis $\alpha_2^{*} = 0$ represents the locus of stable periodic particle orbits which are traversed with a mean angular frequency differing but slightly from ω_1. The stable periodic segments along which $10\sqrt{3\alpha_2^{*}} >> 10\sqrt{\alpha_1^{*}}$ marks those particle orbits which are traversed with a mean frequency close to ω_2.

Curves of D_1 = constant, intersecting all the above curves are also displayed for a few selected values of D_1.

4. FREQUENCIES OF THE PERIODIC MOTIONS

In the present nonlinear treatment, except for the special periodic motions mentioned above which are described only by one single normal mode, all the other periodic particle orbits are generated by a superposition of both normal modes. Periodicity here is achieved as the result of an adjustment of the natural frequencies via the nonlinear

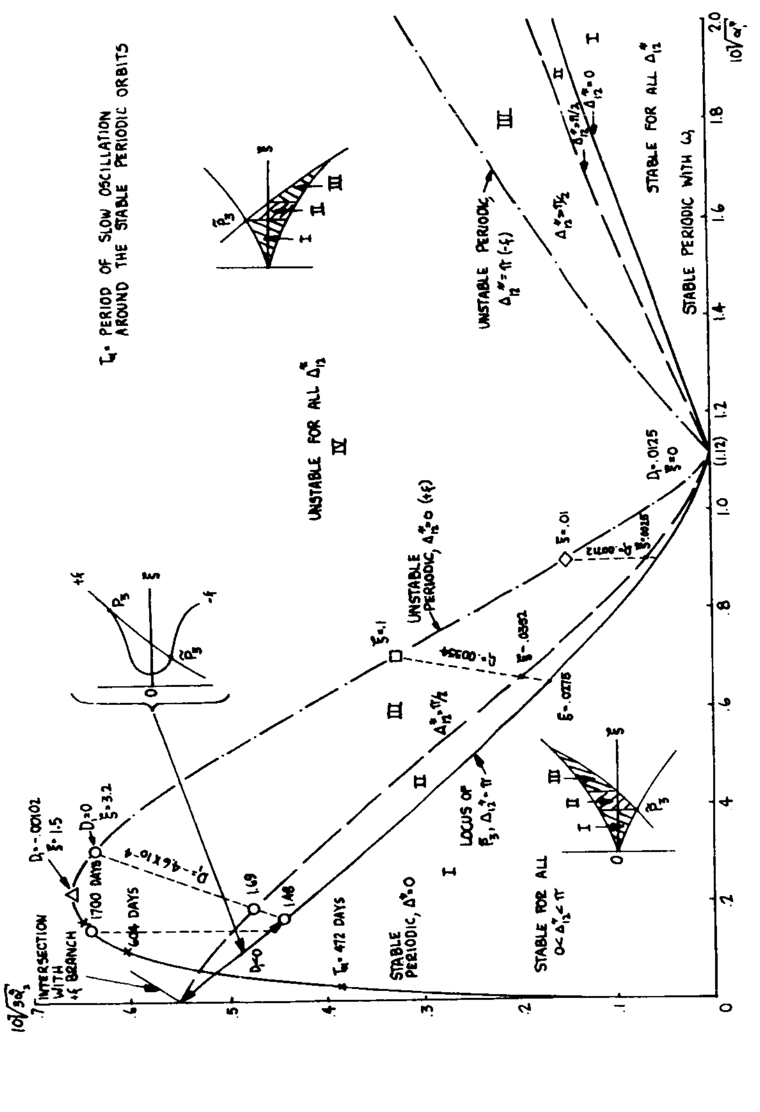


Fig. 7: Tangency Locus of Equilibrium Values for α_1^* and α_2^*

coupling which occurs between the two modes and which makes them exactly commensurable. The resultant frequency shifts $\Delta\omega_1$ and $\Delta\omega_2$ in the original undisturbed frequencies ω_1 and ω_2, lead to normal modes with modified commensurable (3:1) frequencies ω_1' and ω_2'

$$\omega_1' = \omega_1 + \Delta\omega_1 = 3\omega_2' = 3(\omega_2 + \Delta\omega_2) \tag{75}$$

This point was also raised earlier in the discussion following Eq. (72).

The orbital period T is determined by the slower mode

$$T = \frac{2\pi}{\omega_2'} \tag{76}$$

During this time T three cycles of the faster mode are completed.

E. Evaluation of the Frequency Shifts for Periodicity

For every point on the "periodic motion" curve of Fig. 7 there exists a unique set of equilibrium values α_{1E}^* and α_{2E}^*.

The shifts $\Delta\omega_1$ and $\Delta\omega_2$ can be estimated by writing

$$\omega_1(1 + \dot{\beta}_1')t = 3\omega_2(1 - \dot{\beta}_2')t \tag{77}$$

and solving for $\omega_1\dot{\beta}_1'$ and $\omega_2\dot{\beta}_2'$ from the relations

$$\begin{aligned} \dot{\beta}_1^* &= \frac{\partial K_{12}^*}{\partial \alpha_1'} = .02939 + \omega_1\dot{\beta}_1' \\ \dot{\beta}_2^* &= \frac{\partial K_{12}^*}{\partial \alpha_2^*} = .03146 + 3\omega_2\dot{\beta}_2' \end{aligned} \tag{78}$$

evaluated at α_{1E}^* and α_{2E}^*. From here one finds

$$\Delta\omega_1 = \omega_1 \dot{\beta}_1' = .2308\alpha^*_{1E} - 5.1\alpha^*_{2E} \mp \alpha^{*-1/2}_{1E}\alpha^{*3/2}_{2E}$$
$$\Delta\omega_2 = -\omega_2 \dot{\beta}_2' = 1.7\alpha^*_{1E} - 2.039\alpha^*_{2E} \pm \alpha^{*1/2}_{1E}\alpha^{*1/2}_{2E} \qquad (79)$$

where the upper sign corresponds to $\Delta^*_{12} = 0$ and the lower to $\Delta^*_{12} = \pi$.

F. Variation of α^*'s Near Equilibrium Points

For small disturbances from the equilibrium points P_2 and P_3 the time dependence of ξ can be approximated by means of a Taylor series expansion of f and η around the equilibrium points.

Letting

$$f = f_E + (\xi - \xi_E)f_E' + \frac{1}{2!}(\xi - \xi_E)^2 f_E'' + \dots$$

and

$$\eta = \eta_E + (\xi - \xi_E)\eta_E' + \frac{1}{2!}(\xi - \xi_E)^2 \eta_E'' + \dots \qquad (80)$$

and recalling that $\eta_E = f_E$, $\eta_E' = f_E'$, we can combine Eqs. (68) and (80) to obtain (after approximating 23.97 by 24 for convenience)

$$\frac{\dot{\xi}}{24|D_1|} = \frac{1}{24|D_1|}\frac{d}{dt}(\xi - \xi_E) = \sqrt{f^2 - \eta^2} \cong \sqrt{|f_E|(f_E'' - \eta_E'')}\ (\xi - \xi_E)$$

(83)

whence

$$\xi - \xi_E = e^{24|D_1|\sqrt{|f_E|(f_E'' - \eta_E'')}\ t} \qquad (84)$$

For a stable point such as P_2 in Fig. 6(b), we have $f_E'' < \eta_E''$. This makes the exponent in Eq. (84) imaginary of the form $i\omega_\alpha t$ and indicates a slow oscillatory variation in α. For an unstable periodic point such as P_3, $f_E'' > \eta_E''$ which leads to an exponential growth of α with time.

A few representative values of the period $T_\alpha = 2\pi/\omega_\alpha$ are indicated alongside the stable periodic segment in Fig. 7.

The developments of the present section can now be summarized by means of the following general conclusions:

1. On the assumption that the out of plane terms do not couple strongly with the in-plane terms (which will be proven later) it is possible to reduce the problem to an essentially 2-dimensional one.
2. Initial conditions which lie on an η curve located to the left of the limiting curve of type η_3 will lead to bounded motions of the particle in the xy plane.
3. Depending on whether the η curve is tangent to the f curve at a point such as P_3 or P_2, periodic particle motions of an unstable or a stable type, respectively, may exist.
4. The periodic orbits generally result from a superposition of the two normal modes of vibration in which the nonlinear coupling has brought about commensurability of the basic frequencies by means of appropriate frequency shifts. For special initial conditions, periodic particle motions consisting of only the faster normal mode may exist.
5. In the neighborhood of stable equilibrium points of type P_2, the momenta α_1^* and α_2^* perform low frequency bounded oscillations in time. Near unstable equilibrium points of type P_3, the α^*'s will tend to grow exponentially with time, which results in a large growth of the particle's motion in the physical xy plane.

XI. ANALYSIS OF THE INTERNAL OUT OF PLANE MOTION

The analysis of the out of plane motion is rather simple and straightforward compared to the coplanar analysis of Section X. We shall investigate the coupling of α_3^* and α_1^* in the region where $\alpha_2^* = 0$, by neglecting the α_2^* terms in Eq. (61).

The reason for this particular decision is the result of hindsight, based on a prior preliminary study of the external effects on the coplanar motion which disclosed the presence of a stable equilibrium point $\alpha_1^* \neq 0$, $\alpha_2^* = 0$, for the Sun perturbed problem. This will be discussed in more detail in Section XII.

Let us denote by F^* the internal terms left in the Hamiltonian K^* of Eq. (61) when all α_2^* terms are dropped. We have then

$$\begin{aligned} F^* = {} & .1154\alpha_1^{*2} + .09035\alpha_1^*\alpha_3^* C_{2(\beta_1^*-\beta_3^*)} + .08893\alpha_1^*\alpha_3^* \\ & - .002231\alpha_3^{*2} + .074801\alpha_3^* + .02939\alpha_1^* \end{aligned} \tag{85}$$

Let $\Delta_{13}^* = \beta_1^* - \beta_3^*$.

From Hamilton's equations we then obtain

$$\begin{aligned} \dot{\alpha}_1^* &= 2\cdot .09035\alpha_1^*\alpha_3^* S_{2\Delta_{13}^*} \\ \dot{\alpha}_3^* &= -\, 2\cdot .09035\alpha_1^*\alpha_3^* S_{2\Delta_{13}^*} \end{aligned} \tag{86}$$

This leads to the new integral of motion

$$\alpha_1^* + \alpha_3^* = D_2 \tag{87}$$

with $D_2 > 0$.

As we did before for $\dot{\alpha}_2^*$, we can now write for $\dot{\alpha}_1^{*2}$

$$\left(\dot{\alpha}_1^*\right)^2 = 4\cdot(.09035)^2\alpha_1^{*2}\alpha_3^{*2} - 4\Big[F^* - .1154\alpha_1^{*2} - .08893\alpha_1^*\alpha_3^* + .002231\alpha_3^{*2} - .02939\alpha_1^* - .074001\alpha_3^*\Big]^2 \tag{88}$$

We introduce the auxilliary variable ξ_3

$$\xi_3 = \frac{\alpha_3^*}{D_2} \tag{89}$$

and end up again with the equation

$$\left(\frac{\dot{\xi}_3}{.1807D_2}\right)^2 = f^2 - \eta^2 \tag{90}$$

where this time

$$f = \pm\,\xi_3(1 - \xi_3) \tag{91}$$

and

$$f'(o) = \pm\, 1 \tag{92}$$

$$\eta = \frac{F^*}{.09035D_2^2} - 1.277(1 - \xi_3)^2 - .9843\xi_3(1 - \xi_3) + .02469\xi_3^2 - \frac{.32534}{D_2} - \frac{.50259}{D_2}\,\xi_3 \tag{93}$$

At the origin, the first and second η derivatives are

$$\eta'(o) = 1.5703 - \frac{.50259}{D_2} \gtrless 0 \quad \text{depending on } D_2 \tag{94}$$

$$\eta''(o) = \eta''(\xi_3) = -.5366 < 0 \tag{95}$$

$$\eta'(o) < 0 \quad \text{if} \quad D_2 < .3201 \tag{96}$$

and

$$\eta'(o) \le -1 \quad \text{if} \quad D_2 \le .1955 \tag{97}$$

The magnitude of the slope η' will determine the time history of Δ_{13}^*. In particular, if $\eta' < -1$ then Δ_{13}^* will exhibit a circulatory behavior, while a value of $-1 < \eta' < 0$ would lead to a librational behavior.

An upper bound on η' can be established by making a reasonable estimate for an upper value of D_2. Such an estimate can be furnished from some of the mathematical and physical considerations which underlie the present analysis.

From a mathematical standpoint it is clear that in the binominal expansions and truncations used to obtain the expression for the Hamiltonian H(x,y,z,t) of Eq. (21) it was assumed that x,y,z were small compared to unity.

From a physical point of view it is not clear that relatively large displacements away from the Moon would necessarily invalidate the conclusions of the present analysis, but the large accelerations resulting from large displacements towards the Moon or Earth could not be tolerated.

If we assume that the displacements should be limited to values $x,y,z < .5$ (say) then for the excitation mode ω_1 we can obtain from Eqs. (36)

$$\sqrt{\alpha_1^*} < \frac{.5}{2} = .25$$

$$\text{i.e.,} \quad \alpha_1^* < .0625$$

and also

$$\alpha_3^* < .0625$$

so that for these limits

$$D_2 < .1250 < .1955$$

The slopes of all η curves are thus steeper than $f'(\xi_3)$ from which follows that every η curve will intersect both $\pm$ f branches, giving rise to a circulatory motion in Δ_{13}^* as indicated in Fig. 8.

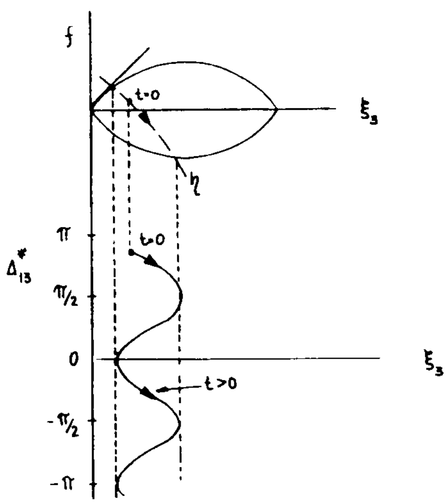


Fig. 8: Geometry in (f, ξ_3) and (Δ_{13}^*, ξ_3) space.

Since $|\eta'|>|f'|$ no equilibrium points with $\alpha_3^* \neq 0$ can exist, and consequently no periodic orbits in xyz space result from the non-linear coupling of modes 1 and 3.

The actual slope of any η curve would depend of course on the value chosen for D_2, subject to the limits mentioned earlier.

We may choose for example a representative value of $\alpha_1^* = .006$ (say) and assume α_3^* to be of the same magnitude (this α_1^* is very close to the actual coplanar equilibrium value of α_1^* in the externally perturbed case discussed in Section XII). Then we have

$$D_2 \simeq 2\alpha_3^* = .012 \tag{98}$$

This results in a slope

$$\begin{gathered} \eta' \simeq 1.57 - \frac{.5026}{.012} = -40.3 = \tan^{-1}\theta \\ \text{or} \\ \theta \simeq 90^{\circ} \end{gathered} \tag{99}$$

In other words the η curve intersects the ξ_3 axis nearly vertically, from which one concludes that $\xi_3 \simeq$ constant; thus, there is hardly any energy interchange taking place between α_3^* and α_1^*, which shows that the out of plane coupling is not very important in this problem, and that the motion is dominated by the coplanar coupling.

That the out of plane coupling does not introduce any instabilities when $\alpha_3^* \ll$ and α_1^* is close to its equilibrium value $\alpha_1^* \simeq .006$ could also have been deduced directly from the expression for F^* in Eq. (85). For very small α_3^* it is sufficient to consider only the terms linear in α_3^*, and to evaluate the coefficients at $\alpha_1^* \simeq .006$. The resultant Mathieu type Hamiltonian F^* indicates a parametrically excited motion. Such Hamiltonians are discussed more fully in Appendix F, (in connection with the solar effects on the coplanar motion examined in Section XII) but under the assumption that the values of α_1^* and α_2^* are to remain very small (i.e., coplanar particle motions for very small perturbations from rest at L_4).

If one applies the results of Appendix F to the present situation. and notes that the coefficient of $\alpha_3^* C_{2(\beta_1^* - \beta_3^*)}$ is smaller than that of α_3^* one readily concludes that the parametric resonance present in the out of plane motion does not lead to instability.

XII. ANALYSIS OF EXTERNAL EFFECTS

1. DETERMINATION OF EQUILIBRIUM POINTS

For a complete analysis of the motion in the presence of the external solar effects, one must retain the complete expression for K^* given in Eq. (61).

From the discussion of Section XI it was seen that the α_1^*, α_3^* internal coupling did not lead to any measurable transfer of energy from the out-of-plane mode to the coplanar mode of motion, while from Section X we have established the existence of an appreciable coplanar coupling effect.

The major long term solar effect causes mainly an excitation of the α_1^* mode. The α_3^* mode does not experience any external excitation to the order of magnitude of the terms retained. This latter statement follows from the developments presented in Section XIII.

If a stable motion in the presence of the Sun is possible in which α_1^*, α_2^* and α_3^* remain small, it would suffice to retain only linear terms in K^* in order to determine long term effects. To linear terms we have the simpler Hamiltonian

$$.02425\alpha_1^* + .02412\alpha_2^* + .07899\alpha_3^* - .02563\alpha_1^* C_{2\beta_1^*} \tag{100}$$

which is of the Mathieu type, as indicated in Appendix F, and leads to parametric resonance in the α_1^* motion.

Since .02563 > .02425, the stability criteria of Appendix F indicate that the motion falls into the unstable region of the Mathieu plane, and that therefore to linear terms no motion can exist for which α_1^* remains very small.

From a physical point of view this means that the libration point L_4 is not stable with respect to small perturbations, when the solar force field is included, and that the higher order terms in K_2^* must be retained in any analysis.

The lack of stability exhibited by the linearized Hamiltonian does not preclude the existence of equilibrium points in the α^* space for the complete Hamiltonian. In view of the negligible effect of α_3^* on the coplanar motion, it is of interest to look for equilibrium points for $\alpha_3^* = 0$. Such points in the (α_1^*, α_2^*) plane are determined by looking for solutions to Hamilton's equations of the form $\dot{\alpha}_1^* = \dot{\alpha}_2^* = 0$.

Once such points are located, it is then necessary to investigate the type of equilibrium which exists there, and to identify the stable ones.

This search is more easily carried out if one switches over to a set of normal canonical coordinates (Q^*, P^*) defined by

$$\begin{pmatrix} Q_1^* \\ Q_2^* \\ P_1^* \\ P_2^* \end{pmatrix} = \begin{pmatrix} \sqrt{2\alpha_1^*} & 0 & 0 & 0 \\ 0 & \sqrt{2\alpha_2^*} & 0 & 0 \\ 0 & 0 & \sqrt{2\alpha_1^*} & 0 \\ 0 & 0 & 0 & \sqrt{2\alpha_2^*} \end{pmatrix} \begin{pmatrix} S_{\beta_1^*} \\ S_{\beta_2^*} \\ C_{\beta_1^*} \\ C_{\beta_2^*} \end{pmatrix} \tag{101}$$

After setting $\alpha_3^* = 0$, the two dimensional part of K^*, which we denote here by K_2^*, becomes

$$\begin{aligned} K_2^* = {} & \frac{.1154}{4}\left(P_1^{*2} + Q_1^{*2}\right)^2 - \frac{5.1}{4}\left(P_1^{*2} + Q_1^{*2}\right)\left(P_2^{*2} + Q_2^{*2}\right) + \frac{3.059}{4}\left(P_2^{*2} + Q_2^{*2}\right)^2 \\ & - \frac{23.97}{4}\left(P_1^*P_2^* - Q_1^*Q_2^*\right)\left(P_2^{*2} + Q_2^{*2}\right) + \frac{.02425}{2}\left(P_1^{*2} + Q_1^{*2}\right) \\ & + \frac{.02412}{2}\left(P_2^{*2} + Q_2^{*2}\right) - \frac{.02563}{2}\left(P_1^{*2} - Q_1^{*2}\right) \end{aligned} \tag{102}$$

The equilibrium points (Q_e^*, P_e^*) are obtained from the solution of the equations

$$\begin{pmatrix} \dot{Q}_e^* \\ \dot{P}_e^* \end{pmatrix} = \Phi_o \begin{pmatrix} K^*_{2P^*T} \\ K^*_{2Q^*T} \end{pmatrix} = 0 \tag{103}$$

From Eq. (103) we have

$$K^*_{2P_1^*} = 0 = .1154P_1^*\left(P_1^{*2} + Q_1^{*2}\right) - \frac{5.1}{2} P_1^*\left(P_2^{*2} + Q_2^{*2}\right)$$

$$- \frac{23.97}{4} P_2^*\left(P_2^{*2} + Q_2^{*2}\right) - .001379P_1^* \tag{104a}$$

$$K^*_{2P_2^*} = 0 = - \frac{5.1}{2} P_2^*\left(P_1^{*2} + Q_1^{*2}\right) + 3.059P_2^*\left(P_2^{*2} + Q_2^{*2}\right)$$

$$- \frac{23.97}{4} \left(3P_1^*P_2^{*2} + P_1^*Q_2^{*2} - 2P_2^*Q_1^*Q_2^*\right) + .02412P_2^* \tag{104b}$$

$$K^*_{2Q_1^*} = 0 = .1154Q_1^*\left(P_1^{*2} + Q_1^{*2}\right) - \frac{5.1}{2} Q_1^*\left(P_2^{*2} + Q_2^{*2}\right)$$

$$+ \frac{23.97}{4} Q_2^*\left(P_2^* + Q_2^*\right) + .04988Q_1^* \tag{104c}$$

$$K^*_{2Q_2^*} = 0 = - \frac{5.1}{2} Q_2^*\left(P_1^{*2} + Q_1^{*2}\right) + 3.059Q_2^*\left(P_2^{*2} + Q_2^{*2}\right)$$

$$- \frac{23.97}{4} \left(2Q_2^*P_1^*P_2^* - Q_1^*P_2^{*2} - 3Q_1^*Q_2^{*2}\right) + .02412Q_2^* \tag{104d}$$

Equations (104c) and (104d) are identically satisfied if we chose $Q^*_{1e} = Q^*_{2e} = 0$. For convenience we shall therefore restrict our search to those equilibrium points for which

$$Q^*_{1e} = Q^*_{2e} = 0 \tag{105}$$

For the above Q^*'s Eqs. (104a) and (104b) give

$$.1154P_1^{*3} - 2.55P_1^*P_2^{*2} - 5.810P_2^{*3} - .001379P_1^* = 0 \tag{106a}$$

$$-2.55P_1^{*2}P_2^* + 3.059P_2^{*3} - 17.43P_1^*P_2^{*2} + .02412P_2^* = 0 \tag{106b}$$

One equilibrium point can be obtained by setting $P_{2e}^* = 0$ (which automatically satisfies Eq. (106b) and then solving for P_{1e}^* from the relation

$$.1154P_1^{*2} - .001379 = 0 \tag{107}$$

or

which corresponds to

$$\begin{aligned} P_{1e}^* &= .1093 \\ \alpha_{1e}^* &= .005975 \end{aligned} \tag{108}$$

The above value of α_1^* is the one which was used in earlier sections when representative numerical values were used.

The first equilibrium point, which we denote by E_I, is thus specified by the coordinates

$$\begin{aligned} E_I: \quad & Q_1^* = Q_2^* = Q_3^* = P_2^* = P_3^* = 0 \qquad & \alpha_1^* &= .005975 \\ & P_1^* = .1093 & \alpha_2^* &= 0 \\ & & \alpha_3^* &= 0 \end{aligned} \tag{109}$$

Another equilibrium point can be found for which $P_2^* \neq 0$, all other homogeneous coordinates remaining the same as for point E_I. The values of P_1^* and P_2^* result from the solution of the algebraic equations (106a) and (106b), after P_2^* is factored out from the latter. The coordinates of the second equilibrium point E_{II} were found to be

$$E_{II}: \quad Q_1^* = Q_2^* = Q_3^* = P_3^* = 0$$

$$P_1^* = .1106 \qquad \alpha_1^* = .006116 \tag{110}$$

$$P_2^* = -.003675 \qquad \alpha_2^* = 6.753 \times 10^{-6}$$

The two points E_I and E_{II} were the only ones readily found for the present simplified conditions. A machine search of the complete set of Eqs. (103) might reveal the existence of additional roots. The periodic elliptic particle motion of mode close to ω_1 corresponding to conditions at E_I has a semimajor axis of about 60,000 mi and a semiminor axis of half this value. These values were determined by computing $r_{max} = [\bar{x}^2 + \bar{y}^2]^{1/2}_{max}$ where the ω_1 modes of $\bar{x}$ and $\bar{y}$ of Eq. (36) were used, and the maximum determined with respect to $\omega_1\beta_1^{\neq}$. It can be shown that this requires that $8.422S_{2\omega_1\beta_1^{\neq}} + 4.423S_{2\omega_1\beta_1^{\neq}+247.14^\circ} = 0$ and results in a value $\omega_1\beta_1^{\neq} \approx 15.62^\circ$. The dimensionless expression for r_{max} then becomes $r_{max} \cong 3.2\alpha_1'^{1/2}$, and at $\alpha_1^* \cong .006$ amounts to roughly $3.2\sqrt{.955:006} \times 2.4 \times 10^5 = 58{,}128 \cong 60{,}000$ in round numbers.

In a similar manner one finds for the maximum dimensionless displacement in mode ω_2 the semimajor axis $r_{max} \cong 9.1\sqrt{\alpha_2'}$ and in miles $r_{max} = 9.1\sqrt{3\omega_2\alpha_2^*} \times 2.4 \times 10^5 = 9.1\sqrt{.8937}/\sqrt{2}\; P^*_{2_{max}} \cdot 2.4 \times 10^5$ miles.

It is of interest to observe that this result indicates the particles mean motion is synchronized with that of the Sun such that their angular positions coincide closely whenever the particle crosses one of the axes of the ellipse.

We recall that at equilibrium $Q_1^* = 0$ and hence $\beta_1^* = n\pi$ with $n = 0, 1 \cdots$. For $n = 0$, Eq. (53) gives

$$\beta_1^* = 0 = .02939t + \omega_1\beta_1' + 14.7 - \epsilon + \epsilon'$$

and from here

$$\omega_1\beta_1^{\neq} = \omega_1 t - .02939t - 14.7 + \epsilon - \epsilon'$$

When the particle crosses the major axis we had $\omega_1 \beta_1^{\neq} = 15.62$, and from the commensurability of angular velocities at E_I, $(\omega_1 - .02939) = 1 - m$. Substitution above gives

$$15.62 + 14.7 = 30.32^{\circ} = (1 - m)t + \epsilon - \epsilon' \equiv \xi$$

as defined by Eq. (B-9). Equation (17) then shows the Sun to be located 30.32° below the x axis, and therefore closely aligned with the major axis of the particle's orbit.

2. STABILITY OF THE EQUILIBRIUM POINTS

The stability of the slow variations around the above periodic equilibrium motions in the xy plane can be determined by setting up the expression for the variation δK^* which results from taking small displacements δQ^* and δP^* around the equilibrium values $Q^*_{ie} = 0$ and P^*_e. Clearly, since E_I and E_{II} are equilibrium points, the coefficients of the linear terms in δP^* must vanish, and on then obtains in three dimensions

$$\begin{aligned}
\delta K^* = {} & .02885P^{*2}_{1e}\left[6\delta P^{*2}_1 + 2\delta Q^{*2}_1 + \cdots\right] - 1.275\left[P^{*2}_{1e}\left(\delta P^{*2}_2 + \delta Q^{*2}_2\right)\right. \\
& \left. + 4P^*_{1e}P^*_{2e}\delta P^*_1\delta P^*_2 + P^{*2}_{2e}\left(\delta P^{*2}_1 + \delta Q^{*2}_1\right) + \cdots\right] \\
& + .7648P^{*2}_{2e}\left[6\delta P^{*2}_2 + 2\delta Q^{*2}_2 + \cdots\right] - .00039\left[P^{*2}_{1e}\left(\delta P^{*2}_3 + \delta Q^{*2}_3\right)\right. \\
& \left. + 4P^*_{1e}P^*_{3e}\delta P^*_1\delta P^*_3 + P^{*2}_{3e}\left(\delta P^{*2}_1 + \delta Q^{*2}_1\right)\right] \\
& + .04958\left[P^{*2}_{1e}\delta P^{*2}_3 + P^{*2}_{3e}\delta P^{*2}_1 + 2P^*_{1e}P^*_{3e}\left(2\delta P^*_1\delta P^*_3 + \delta Q^*_1\delta Q^*_3\right)\right] \\
& + .2113\left[P^{*2}_{2e}\left(\delta P^{*2}_3 + \delta Q^{*2}_3\right) + 4P^*_{2e}P^*_{3e}\delta P^*_2\delta P^*_3\right. \\
& \quad \left. + P^{*2}_{3e}\left(\delta P^{*2}_2 + \delta Q^{*2}_2\right)\right] - .0005578\left[P^{*2}_{3e}\left(6\delta P^{*2}_3 + 2\delta Q^{*2}_3\right)\right] \\
& \quad + .0395\left[\delta P^{*2}_3 + \delta Q^{*2}_3\right] - 5.810\left[P^*_{1e}P^*_{2e}\left(3\delta P^{*2}_2 + \delta Q^{*2}_2\right)\right.
\end{aligned}$$

(con't on next page)

$$+ P_{2e}^{*2}\left(3\delta P_1^*\delta P_2^* - \delta Q_1^*\delta Q_2^*\right) + .012126\left[\delta P_1^{*2} + \delta Q_1^{*2}\right]$$

$$+ .012062\left[\delta P_2^{*2} + \delta Q_2^{*2}\right] - .012815\left[\delta P_1^{*2} - \delta Q_1^{*2}\right] \tag{111}$$

Applying expression (111) to point E_I results in

$$\delta K^* = .001380\delta P_1^{*2} + .025638\delta Q_1^{*2} - .003174\delta P_2^{*2} - .003174\delta Q_2^{*2}$$

$$+ .04008\delta P_3^{*2} + .03949\delta Q_3^{*2} \tag{112}$$

Since for every value of i = 1,2,3 the coefficients of δP_i^* have the same sign as the coefficients of δQ_i^{*2} (i.e., δK^* is either positive or negative definite irrespective of the signs of δP^* or δQ^*) we can conclude that point E_I is stable for small disturbances in all principal directions. The period of the slow variations in $\delta P_1, \delta Q_1$ is approximately 83 months.

It is more convenient to retain only coplanar terms in δK^* for the determination of stability at E_{II}. We then obtain the expression

$$\delta K^* = .001411\delta P_1^{*2} + .025638\delta Q_1^{*2} + .001838\delta P_1^*\delta P_2^* + .003652\delta P_2^{*2}$$

$$+ 7.847 \times 10^{-5}\ \delta Q_1^*\delta Q_2^* - .001150\delta Q_2^{*2} \tag{113}$$

If we now assume P_1^* and Q_1^* to remain unchanged while we introduce variations δP_1^* and δQ_1^* we have

$$\delta K^* = .001411\delta P_1^{*2} + .025638\delta Q_1^{*2} \tag{114}$$

where $\delta Q_2^* = \delta P_2^* = 0$.

Thus δK^* is positive definite for variations in the first set of coordinates and hence δQ_1^* and δP_1^* remain bounded.

Repeating the same steps for δP_2^* and δQ_2^* while keeping P_1^* and Q_1^* fixed gives

$$\delta K^* = .003652\delta P_2^{*2} - .001150\delta Q_2^{*2} \tag{115}$$

where $\delta P_1^* = \delta Q_1^* = 0$.

Since δK^* is not definite for arbitrary choices of δP_2^* and δQ_2^* we conclude that point E_{II} is not stable in δP_2^* and δQ_2^*, and hence is an unstable equilibrium point. The equilibrium for variations δP_3^* and δQ_3^* was found to be stable, which is in agreement with the findings of the last section.

The above conclusion could have been reached also more rigorously in a somewhat lengthier fashion by writing down the complete system of first order linear differential equations for $\delta\dot{Q}^*$ and $\delta\dot{P}^*$ obtained from δK^* of Eq. (113), and examining the roots of the appropriate characteristic equation. We would find that

$$\begin{pmatrix} \delta\dot{Q}_1^* \\ \delta\dot{Q}_2^* \\ \delta\dot{P}_1^* \\ \delta\dot{P}_2^* \end{pmatrix} = \begin{pmatrix} 0 & 0 & .002822 & .001838 \\ 0 & 0 & .001838 & .007304 \\ -.05126 & -7.85\cdot10^{-5} & 0 & 0 \\ -7.85\cdot10^{-5} & -.0023 & 0 & 0 \end{pmatrix} \begin{pmatrix} \delta Q_1^* \\ \delta Q_2^* \\ \delta P_1^* \\ \delta P_2^* \end{pmatrix} \tag{116}$$

A trial solution of the form e^{st} would lead to the characteristic equation

$$s^4 + 1.282\cdot10^{-4}s^2 - 2.031\cdot10^{-8} = 0 \tag{117}$$

which has one positive root because of the negative constant term. Equation (117) thus bears out the conclusions reached from Eq. (115).

A simple geometrical description of the stable and unstable regions in the 6 dimensional P^*,Q^* space is of course not feasible. On the other hand it is possible to take advantage of the fact that the stable point E_I is noticed to lie very close to the unstable point E_{II}.

It is thus of particular interest to determine the extent of the stable region around E_I, by expanding K^* up to cubic powers in δP^* and δQ^* around E_I.

The intersection of surfaces of constant K^* with the (P_2^*, Q_2^*) plane, for a value of $P_1^* = .11$, is shown in Fig. (9). The dashed curve shows the separatrix which passes through E_{II} and separates the stable from the unstable regions.

In the physical xy plane, a point in the stable region gives rise to slow variations of the elements of the periodic particle orbit corresponding to E_I. A point in the unstable region of the (P_2^*, Q_2^*) plane would lead to large particle departures from the equilibrium orbit, and thus indicate a possible divergence.

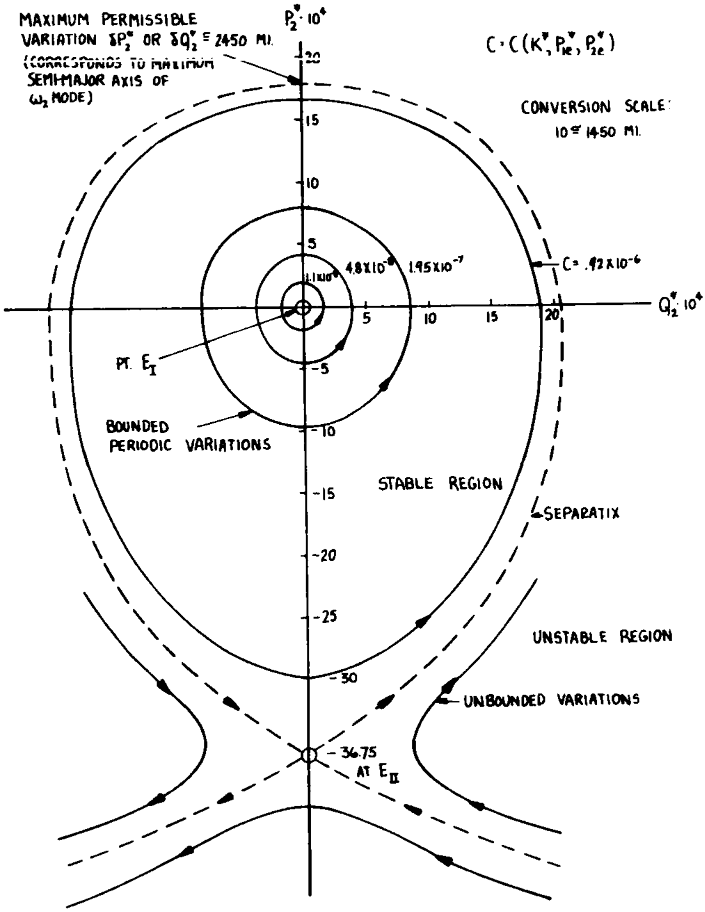


Fig. 9: Stability Regions in the (P_2^*, Q_2^*) Plane Near The Coplanar Equilibrium Points.

XIII. EVALUATION OF THE EFFECT OF THE RESONANCE CAUSED BY THE FORCED SOLUTION $\tilde{z}$

We had alluded on page 23 to the fact that no forced solutions in z, i.e., $\tilde{z}$, had been retained since they are of $o(m^3)$ and would thus give rise to terms of $o(m^5)$ or higher in H when one went on to derive the long period contributions.

A closer second look at the external z terms in $H^{(o)}$ disclosed the existence of a very closely tuned forcing term in the linearized out of plane z motion which could introduce perhaps small divisors in the solution for $\tilde{z}$ and thus depress the order of magnitude of that solution. This would introduce another important long period term into the Hamiltonian K. The resonance in question arises for example from a term such as

$$x_s z_s = -r_{13}^2 \cdot \frac{1}{2} \sin i_o \sin [1.0040212t + \epsilon]$$

which would lead to a detuning of magnitude

$$1.0040212 - 1 = .0040212 \tag{118}$$

This value would introduce a much slower term in K than any of the terms previously retained, and might conceivably require a redefinition of the angular variable β_3^* introduced earlier.

The developments indicated briefly below disclosed that the z resonance terms cancel each other exactly, and consequently do not contribute a term slower than the one already considered. No further modifications to the analysis of the out-of-plane motion of Section XI were thus required. The steps leading to the above mentioned cancellation were nevertheless found interesting enough to justify their inclusion here.

The z portion of the external part of $H^{(o)}$ of Eq. (21) was

$$\overset{(o)}{H(z)} = -m^2\left[\frac{3}{2r_{13}^2}\left(x_s + \sqrt{3}y_s\right)z_s z\right] + \frac{1}{2}\left(\upsilon_x + \sqrt{3}\upsilon_y\right)z + \frac{1}{2}\left(\upsilon_y - \sqrt{3}\upsilon_x\right)P_z \tag{119}$$

For a coordinate system with its x axis pointing at the instantaneous position of the Moon, the angular velocity components υ_x and υ_y are given by

$$\begin{aligned} \upsilon_x &= \dot{i} \cos \eta + \dot{\Omega} \sin i \sin \eta \\ \upsilon_y &= \dot{\Omega} \sin i \cos \eta - \dot{i} \sin \eta \end{aligned} \tag{120}$$

These are the same as Eqs. (B-3) except that η_o has now been replaced by $\eta = gnt + \epsilon - \Omega$ and $g = 1.0040212$.

The angular velocities $\dot{\Omega}$ and $\dot{i}$ can be expressed in terms of η,i and the solar acceleration component W normal to the Earth-Moon plane at the Moon's position, by means of the variational equations on page 404 of Ref. 9, in which a corresponds to $\langle r_{12} \rangle$ here

$$\begin{aligned} \dot{\Omega} &= \frac{r_{12} \sin \eta}{na^2 \sin i} W \\ \dot{i} &= \frac{r_{12} \cos \eta}{na^2} W \end{aligned} \tag{121}$$

By our nondimensionalization convention $na^2 = 1$, so that

$$\upsilon_y = \frac{r_{12}}{D} W\left[\sin \eta \cos \eta - \cos \eta \sin \eta\right] = 0 \tag{122}$$

and the angular velocity $\bar{\omega}$ has thus no component in the y direction.

One can also write for υ_x

$$\upsilon_x = \frac{r_{12}}{Dna^2} W\left[\sin^2\eta + \cos^2\eta\right] = (1 + \rho)W \tag{123}$$

A suitable expression for W can be obtained by differentiation from the potential energy V_s near the Moon. If we let x,y,z denote small displacements from the instantaneous position of the Moon, we have

$$W = - \left.\frac{\partial V_s}{\partial z}\right|_{x,y,z=0} = \frac{\partial}{\partial z}\left\{ m^2 \left[\frac{3}{2r_{13}^2} \left(\bar{r}_{13} \cdot \bar{r}_{14}\right)^2 - \frac{1}{2} r_{14}^2 \right] \right\}_{x,y,z=0}$$

$$= \frac{\partial}{\partial z}\left\{ m^2 \left[\frac{3}{2r_{13}^2} \left(X_s(1 + \rho + x) + y_s y + z_s z\right)^2 - \frac{1}{2}\left((1 + \rho + x)^2 + y^2 + z^2\right)\right]\right\}_{x,y,z=0}$$

$$= 3m^2 \frac{x_s z_s}{r_{13}^2} + o(m^5) = 3m^2 \sin i \cos \xi \sin (\Omega - v') \qquad (124)$$

To sufficient accuracy then

$$\upsilon_x \cong W = 3m^2 \sin i_o \cos \xi \sin (\Omega - v') \qquad (125)$$

To check if $\overset{(o)}{H}(z)$ would in fact lead to the presence of small divisors in the solution for $\tilde{z}$ it suffices to check if $\overset{(o)}{H}(z)$ contains slowly varying terms of frequency .0040212 when we replace in it z and P_z by the homogeneous solutions $\bar{z}$ and $\bar{P}_z$.

$$\overset{(o)}{H}(z) = \frac{3}{4} m^2 \sin i_o \sqrt{2\alpha_3} \cos \beta_3^{\neq} \left\{\sin (1.0040212t + \epsilon) + \sqrt{3} \cos (1.0040212t + \epsilon)\right\} + \frac{1}{2} \upsilon_x \sqrt{2\alpha_3} \left[\cos \beta_3^{\neq} + \sqrt{3} \sin \beta_3^{\neq}\right]$$

(126)

Since

$$\Omega - \bar{v}' - \xi = -1.0040212t - \epsilon \tag{127}$$

we may retain in υ_x only the dominant resonance term

$$\upsilon_x \cong -\frac{3}{2} m^2 \sin i_o \sin (1.0040212t + \epsilon) \tag{128}$$

When Eq. (128) is substituted into $H^{(o)}(z)$ of Eq. (126) and all the terms combined it is found that all the long period terms cancel each other exactly and only fast terms remain. From this one can conclude that the forcing function of the linearized z equation does not contain a resonance term which is close enough to introduce small divisors into the forced response $\tilde{z}$ and thereby lower its order of magnitude from $o(m^3)$ to $o(m^2)$ or less.

Based on the foregoing we can conclude that the neglect of the contribution of $\tilde{z}$ to the long period terms $\overline{(\partial H/\partial z)\tilde{z}}$ was consistent with our convention of neglecting terms of order higher than $o(m^4)$.

This analysis shows that although the Sun has an appreciable long term effect on the changes in inclination of the lunar orbital plane, it has the same effect also on the orbital plane of the librating particle, with the net result that any relative long term out-of-plane responses vanish. Short period, fast, relative terms do not cancel out though.

XIV. SUMMARY AND CONCLUSIONS

In the present dissertation, the 3-dimensional stability of the motion of a particle near the equilateral libration points of the Earth-Moon system, in the presence of the Sun, has been investigated.

Because the inclusion of lunar eccentricity would have introduced into the problem a larger number of internal and external resonances than could have been handled by the present method of approach, it was found necessary to restrict the stability analysis to a lunar orbit perturbed by the Sun but without eccentricity.

Four major conclusions emerge from the present study. First, small coplanar motions near L_4 or L_5 will grow large because of parametric excitation by the Sun, as a result of nonlinear resonance. In fact, the growth of the energy in the faster normal mode of the linearized theory is found to be governed by a Mathieu equation.

Second, the out-of-plane motion is not seriously excited by the Sun, and has a negligible effect on the coplanar motion, which is the dominant factor as far as stability is concerned.

Third, a stable periodic coplanar orbit can exist in the presence of the Sun. It consists of a clockwise motion along the 1:2 ellipse corresponding to the first (or faster) normal mode, and has a semimajor axis of approximately 60,000 mi. The external nonlinear excitation causes the mean angular motion of the particle to become synchronized with that of the Sun. Thus to an observer located at L_4 and looking continuously in the direction of the Sun, the particle would appear to move back and forth across his line of sight in the manner of a simple harmonic oscillator. The times of crossing of the line of sight coincide closely with the times at which the line of sight is aligned with the major or minor axis of the ellipse.

Fourth, the presence of the internal resonant excitation, resulting from the near commensurability (3:1) of the two coplanar normal models makes the stability somewhat delicate. As a consequence, the semimajor axis of the second mode is limited to magnitudes less than approximately 2400 mi. For larger values the motion becomes unstable and may result in very large displacements which would exceed the range of applicability of the present theory.

Appendix A

SOLAR GRAVITATIONAL GRADIENT CONTRIBUTION

Consider the term

$$\frac{1}{r_{34}} - \frac{\bar{r}_{13} \cdot \bar{r}_{14}}{r_{13}^3} \tag{A-1}$$

of Eq. (7), and decompose $\bar{r}_{34}$ into

$$\bar{r}_{34} = \bar{r}_{31} + \bar{r}_{14} = \bar{A} \text{ (for simplicity)} \tag{A-2}$$

For $(r_{14}/r_{31}) \ll 1$ we can expand $1/r_{34}$ into a Taylor series around r_{31} as shown:

$$\frac{1}{r_{34}} = \frac{1}{\left[\bar{r}_{34} \cdot \bar{r}_{34}\right]^{1/2}} = \frac{1}{r_{31}} + \bar{r}_{14} \cdot \nabla \left.\frac{1}{\left[\bar{A} \cdot \bar{A}\right]^{1/2}}\right|_{\bar{r}_{14}=o}$$
$$+ \frac{1}{2}\left(\bar{r}_{14} \cdot \bar{r}_{14}\right) \nabla^2 \left.\frac{1}{\left[\bar{A} \cdot \bar{A}\right]^{1/2}}\right|_{\bar{r}_{14}=o} + 0\left(\frac{r_{14}}{r_{13}}\right)^3 + \cdots \tag{A-3}$$

where

$$\nabla = \frac{\partial}{\partial \bar{r}_{31}} \quad \text{and} \quad \bar{r}_{31} = - \bar{r}_{13}$$
$$\nabla \frac{1}{\left[\bar{A} \cdot \bar{A}\right]^{1/2}} = - \frac{\bar{A} \cdot \nabla\bar{A}}{\left[\bar{A} \cdot \bar{A}\right]^{3/2}} = - \frac{\dot{A} \cdot \underline{\underline{I}}}{\left[\bar{A} \cdot \bar{A}\right]^{3/2}} \tag{A-4}$$

and

$$\bar{r}_{14} \cdot \nabla \frac{1}{\left[\bar{A} \cdot \bar{A}\right]^{1/2}}\Bigg|_{\bar{r}_{14}=o} = \frac{\bar{r}_{13} \cdot \bar{r}_{14}}{r_{13}^3} \qquad \text{(A-5)}$$

where

$\underline{\underline{I}}$ = unit diadic

Similarly

$$\frac{1}{2}\left(\bar{r}_{14} \cdot \bar{r}_{14}\right) \nabla^2 \frac{1}{\left[\bar{A} \cdot \bar{A}\right]^{1/2}} \rightarrow \frac{1}{2}\,\bar{r}_{14} \cdot \nabla \left(- \frac{\bar{r}_{31}}{r_{31}^3}\right) \cdot \bar{r}_{14}$$

$$= \frac{1}{2}\,\bar{r}_{14} \cdot \left[\frac{3}{r_{31}^4}\,\bar{r}_{31}\bar{r}_{31} - \frac{\underline{\underline{I}}}{r_{31}^3} \cdot \bar{r}_{14}\right] = \frac{1}{r_{31}^3}\left[\frac{3}{2}\left(\frac{\bar{r}_{14} \cdot \bar{r}_{31}}{r_{31}}\right)^2 - \frac{1}{2}\;\bar{r}_{14} \cdot \bar{r}_{14}\right]$$

$$= \frac{1}{r_{13}^3}\left[\frac{3}{2}\left(\frac{\bar{r}_{13} \cdot \bar{r}_{14}}{r_{13}}\right)^2 - \frac{1}{2}\,\bar{r}_{14} \cdot \bar{r}_{14}\right] \qquad \text{(A-6)}$$

Combining (A-1), (A-3), (A-5) and (A-6) and neglecting the first term of the series, $1/r_{13}$, which makes no contribution to the equations of motion, we end up with the last term of Eq. (8) which is the expression of (A-6), and represents the solar gradient force near the Earth.

Appendix B

THE EXPRESSIONS FOR $\rho(t)$ AND $\nu(t)$ FROM LUNAR THEORY

The expression for $\rho(t)$ is readily obtained from Eq. (1), p. 281 of Ref. 7, after computing $(a/r)^{-1} = 1 + \rho(t) = r_{12}$ and retaining only terms of $o(m^2)$ or lower. The term $-.00093$ of our Eq. (15) corresponds to $-\frac{1}{6}m^2$ in the series for $(a/r)^{-1}$. The semimajor axis a is set equal to the reference length D in our notation.

Derivation of the expression for $\nu(t)$ requires a few more algebraic manipulations. We shall make use for this of Fig. 2 (p. 13) and Fig. 4 (p. 38) of Ref. 8, which are combined for convenience in Fig. B-1, and also Fig. B-2 which shows the lunar orbital plane as viewed from above (i.e., looking in the direction of the negative Z axis). In order to facilitate the derivation we shall retain (in this Appendix only) the notation and symbols of Ref. 8 irrespective of the use to which some of the letters have been put in the main body of the present report. Where necessary, the corresponding letters in our notation will be pointed out.

In dimensional symbols we now have

$$\begin{aligned}
\bar{\omega} &= \left(n + \nu_z\right) \bar{i}_{z_e} + \dot{i}\bar{i}_N + \dot{\Omega} \sin i \left(\bar{i}_z \times \bar{i}_N\right) \\
\bar{i}_N &= \cos \eta_o \bar{i}_{x_e} - \sin \eta_o \bar{i}_{y_e} \\
\bar{i}_z \times \bar{i}_N &= \sin \eta_o \bar{i}_{x_e} + \cos \eta_o \bar{i}_{y_e} \\
\eta_o &= nt + \epsilon - \Omega
\end{aligned} \tag{B-1}$$

so that

$$\begin{aligned}
\bar{\omega} = &\left(n + \nu_z\right) \bar{i}_{z_e} + \left[\dot{i} \cos \eta_o + \dot{\Omega} \sin i \sin \eta_o\right] \bar{i}_{x_e} \\
&+ \left[\dot{\Omega} \sin i \cos \eta_o - \dot{i} \sin \eta_o\right] \bar{i}_{y_e}
\end{aligned} \tag{B-2}$$

The dimensionless form of $\bar{\omega}$ results if we set n = 1 in (B-2). Noting that $\bar{i}_{x_e}$, $\bar{i}_{y_e}$, $\bar{i}_{z_e}$ are parallel respective to the unit vectors $\bar{i}_x$, $\bar{i}_y$, $\bar{i}_z$ of our L_4 centered coorindate frame, we have

$$\begin{aligned} \upsilon_x &= \dot{\Omega} \sin i \sin \eta_o + \dot{i} \cos \eta_o \\ \upsilon_y &= \dot{\Omega} \sin \cos \eta_o - \dot{i} \sin \eta_o \end{aligned} \tag{B-3}$$

both of which are of $o(m^3)$ or higher. The expression for υ_z can be obtained by taking the time derivative of the true anomaly v in either one of the expressions on p. 110 of Ref. 8 or Eq. (2), p. 281 of Ref. 7. This results in the coefficient of $\bar{i}_z$ of our expression (16).

With the aid of Fig. (B-1) it is also relatively straightforward to determine the components of $\bar{r}_{13}$ in the $\bar{i}_x$, $\bar{i}_y$ and $\bar{i}_z$ directions.

We refer the reader to pp. 38, 41, and 79 of Ref. 8 for a more detailed presentation of the relations summarized here. For convenience the following explanatory relations for the various angular arcs are summarized below.

$$\begin{aligned} Ex \text{ or } E\gamma &= \text{fixed reference line in ecliptic} \\ \Omega\, m'(t) &= v' - \Omega \text{ where } x\,\Omega \equiv \vee\,\Omega = \text{arc of nodal regression} \\ \varpi &= x\,\Omega + \Omega\,A \text{ (measured in two planes)} = \gamma EA \\ \epsilon &= \gamma\, EM_o(o) \text{ i.e., at } t = o \\ s &= \tan M'M \\ v &= xM' = \text{ecliptic projection of } xM \\ i &= M'\,\Omega\,M \\ \gamma &= \tan i \simeq \sin i \\ \eta_o &= \Omega\, M_o = nt + \epsilon - \Omega \\ \epsilon' &= \gamma\, Em'(o) \text{ at } t = o \text{ if } e_s \neq o \end{aligned}$$

One can then show that

$$\cos Mm' = \cos (v - v') \cos M'M = \left(1 - \frac{1}{2} s^2 + \frac{3}{8} s^4 - \ldots\right) \cos (v - v') \tag{B-4}$$

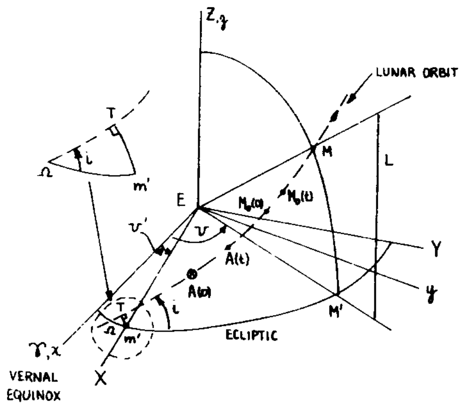
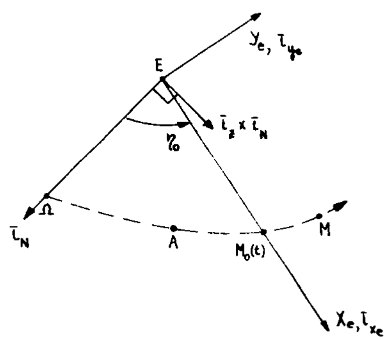


Fig. B-1: Lunar Motion Geometry in 3-Dimension.

m' = SUN
M = MOON
M_o = MEAN LUNAR POSITION
$\upsilon = XM'$ - LUNAR LONGITUDE
υ' = SOLAR LONGITUDE
A = POSITION OF PERILUNE
Ω = DESIGNATES NODAL CROSSING (USED BY US TO DESIGNATE ANGLE $\measuredangle \Upsilon E \Omega$)
$\bar{i}_N$ = UNIT VECTOR FROM E TO Ω



$$\bar{i}_{z_e} = \bar{i}_{x_e} \times \bar{i}_{y_e}$$

Fig. B-2: Planar view of lunar orbital plane and since

$$s = \left(1 - e^2 - \frac{1}{8}\gamma^2\right)\gamma \sin \eta_o + e\gamma \sin \ldots + \frac{1}{8} e^2\gamma \sin \ldots + \ldots \tag{B-5}$$

(from p. 41 of Ref. 8).
while

$$s^2 \leq \gamma^2 \cong \sin^2 i \cong \sin^2 i_o = \sin^2 (5^o 8' 43'') \simeq .008 \ll 1 \tag{B-6}$$

we can approximate to sufficient accuracy

$$\begin{aligned} \cos Mm' &\simeq \cos (v - v') = \cos [nt + \epsilon - n't - \epsilon' \\ &\quad + e, e' \text{ times periodic terms}] \\ &\cong \cos [(1 - m)t + \epsilon - \epsilon'] + \text{higher order terms} \\ &\quad (\text{we have divided by } n = 1) \end{aligned} \tag{B-7}$$

$$\sin Mm' \cong - \sin [(1 - m)t + \epsilon - \epsilon'] + \text{H.O.T.} \tag{B-8}$$

$$\text{Define:} \quad \xi = \sphericalangle Mm' = (1 - m)t + \epsilon - \epsilon' \tag{B-9}$$

and note that

$$\sin m'T = \sin i \sin \Omega m' = \sin i \sin (v' - \Omega) \tag{B-10}$$

With the above relations we can now obtain Eqs. (17) of the text

$$\begin{aligned} x_s &= r_{13} \cos \xi \\ y_s &= - r_{13} \sin \xi \\ z_s &= - r_{13} \sin i \sin (v' - \Omega) = r_{13} \sin i \sin (\Omega - v') \end{aligned} \tag{B-11}$$

Appendix C

TAYLOR SERIES EXPANSION AROUND L_4

The steps needed in the expansion of the various terms in the Lagrangian L of Eq. (8) up to fourth order terms [i.e., $o(m^4)$] are indicated below. In dimensionless notation we have

$$\bar{r}_{14} = \bar{r}_{1L} + \bar{r} \qquad \text{where} \qquad \bar{r}_{1L} = \frac{1}{2}(1+\rho)\,\bar{i}_x + \frac{\sqrt{3}}{2}(1+\rho)\,\bar{i}_y$$

and thus

$$r_{14}^2 = \left[\frac{1}{2}(1+\rho) + x\right]^2 + \left[\frac{\sqrt{3}}{2}(1+\rho) + y\right]^2 + z^2 = \ldots \text{ algebra}$$

$$= 1 + \rho(2 + x + \sqrt{3}\,y) + \left(x + \sqrt{3}\,y + x^2 + y^2 + z^2\right)$$

$$= 1 + (a + b) = 1 + I \qquad \text{(C-1)}$$

where a and b refer to the two terms following 1.

This enables us to write r_{14}^{-1} in the form

$$r_{14}^{-1} = [1 + I]^{-1/2} = 1 - \frac{1}{2} I + \frac{3}{8} I^2 - \frac{15}{48} I^3 + \frac{5}{16} \cdot \frac{7}{8} I^4 + \ldots \qquad \text{(C-2)}$$

A similar expression applies also to r_{24}^{-1} after replacing x by -x in Eq. (C-1).

Evaluate now the various terms in (C-2).

$$\left[I^2\right]: \qquad = \overset{o(m^5)}{\cancel{a^2}} + 2ab + b^2$$

$$2ab = 2\rho(2 + x + \sqrt{3}\,y)\left(x + \sqrt{3}y + x^2 + y^2 + z^2\right)$$

$$= 4\rho\left(x + \sqrt{3}y + x^2 + y^2 + z^2\right) + 2\rho\left(x + \sqrt{3}y\right)^2 + o(m^5)$$

$$= 2\rho\left[2(x + \sqrt{3}y) + 3x^2 + 5y^2 + 2z^2 + 2\sqrt{3}\, xy\right]$$

Terms independent of x, y, or z have been dropped since they don't contribute to the final D.E.

$$b^2 = x^2 + 2\sqrt{3}\, xy + 3y^2 + 2(x + \sqrt{3}y)\left(x^2 + y^2 + z^2\right)$$

$$+ \left(x^2 + y^2\right)^2 + 2\left(x^2 + y^2\right) z^2 + z^4$$

$\left[I^3\right]$: $$= \cancel{a^3} + \cancel{3a^2b} + 3ab^2 + b^3$$

neglect as H.O.T.

$$3ab^2 \rightarrow 6\rho(x + \sqrt{3}y)^2 + o(m^5)$$

$$b^3 \rightarrow (x + \sqrt{3}y)^3 + 3(x + \sqrt{3}y)^2 \left(x^2 + y^2 + z^2\right) + o(m^5)$$

$\left[I^4\right]$: only b^4 contributes

$$b^4 = (x + \sqrt{3}y)^4 + o(m^5)$$

Combining the above terms and neglecting noncontributing factors gives

$$r_{14}^{-1} = -\frac{1}{2}\left[\rho(x + \sqrt{3}y) + \left(x + \sqrt{3}y + x^2 + y^2 + z^2\right)\right]$$

$$+ \frac{3}{8}\left\{2\rho\left[2(x + \sqrt{3}y) + 3x^2 + 5y^2 + 2z^2 + 2\sqrt{3}\, xy\right]\right.$$

$$\left. + \left[(x + \sqrt{3}y)^2 + 2(x + \sqrt{3}y)\left(x^2 + y^2 + z^2\right) + \left(x^2 + y^2\right)^2\right.\right.$$

$$+ 2\left(x^2 + y^2\right) z^2 + z^4\Big\} - \frac{15}{48}\left\{6\rho(x + \sqrt{3}y)^2 + (x + \sqrt{3}y)^3\right.$$

$$\left. + 3(x + \sqrt{3}y)^2 \left(x^2 + y^2 + z^2\right)\right\} + \frac{35}{128}(x + \sqrt{3}y)^4 \qquad \text{(C-3)}$$

Furthermore

$$\bar{r}_{12} \cdot \bar{r}_{14} = (1 + \rho)\left[\frac{1}{2}(1 + \rho) + x\right] = \frac{1}{2}(1 + \rho)^2 + (1 + \rho)\, x$$

$$r_{12}^3 = (1 + \rho)^3 = 1 + 3\rho + \ldots$$

Thus

$$\frac{\bar{r}_{12} \cdot \bar{r}_{14}}{r_{12}^3} = \frac{1}{2}(1 + \rho)^{-1} + (1 + \rho)^{-2} \to x - 2\rho x + \text{noncontributing terms} \qquad \text{(C-4)}$$

The Lagragian L in Eq. (8) is made dimensionless by multiplying it by $D/(\mu_1 + \mu_2)$. Let us multiply Eq. (8) by this factor and then set

$$\mu_1 + \mu_2 = 1$$

$$D = 1$$

and introduce the dimensionless quantity

$$\bar{\mu}_1 = \frac{\mu_1}{\mu_1 + \mu_2} \qquad \text{(C-5)}$$

while from before we had defined already the quantity μ

$$\mu = \frac{\mu_2}{\mu_1 + \mu_2}$$

by Eq. (12).

It then follows that

$$\bar{\mu}_1 + \mu = 1 \tag{C-6}$$

Consider the contribution V_{EM} of Earth and Moon to the potential energy term in L of Eq. (8)

$$V_{EM} = -\left[\frac{\bar{\mu}_1}{r_{14}} + \frac{\mu}{r_{24}}\right] + \mu\left[\frac{\bar{r}_{12} \cdot \bar{r}_{14}}{r_{12}^3}\right] \tag{C-7}$$

We recall that only the x coordinate changes sign when we use the expression for r_{14}^{-1} to obtain r_{24}^{-1}. A convenient expression for V_{EM} can be obtained by making in r_{24}^{-1} the following substitution for all odd powers of x

$$-x^{2n+1} = x^{2n+1} - 2x^{2n+1} \qquad (n = 0,1)$$

When use is made of Eq. (C-6) and the lengthy algebraic manipulations are carried out, one ends up with a V_{EM} given by

$$\begin{aligned} V_{EM} = {} & \left\{\frac{1}{8} x^2 - \frac{5}{8} y^2 + \frac{1}{2} z^2 - \frac{3\sqrt{3}}{4} (1 - 2\mu)\, xy \right. \\ & \left. - \rho(x + \sqrt{3}y)\right\}^{(o)} + \left\{- 7\, \frac{1 - 2\mu}{16}\, x^3 + \frac{3\sqrt{3}}{16}\, y^3 \right. \\ & \left. + \frac{1 - 2\mu}{16} \cdot 33\, xy^2 - \frac{1 - 2\mu}{16} \cdot 12\, xz^2 + \frac{3\sqrt{3}}{16}\, x^2y - \frac{12\sqrt{3}}{16}\, yz^2\right\}_3 \\ & + \left\{\frac{37}{128}\, x^4 - \frac{123}{64}\, x^2y^2 + \frac{3}{16}\, x^2z^2 + \frac{33}{16}\, y^2z^2 - \frac{3}{128}\, y^4 - \frac{3}{8}\, z^4 \right. \\ & \left. + \frac{25\sqrt{3}}{32} (1 - 2\mu)\, x^3y - \frac{45(1 - 2\mu)\sqrt{3}}{32}\, xy^3 + \frac{15(1 - 2\mu)}{8}\, xyz^2\right\}_4 \\ & + \rho \left\{- \frac{3}{8}\, x^2 + \frac{15}{8}\, y^2 - \frac{3}{2}\, z^2 + \frac{9}{4} (1 - 2\mu)\, xy\right\}_s \end{aligned} \tag{C-8}$$

The solar contribution V_s to the potential energy term in L is found below.

$$\bar{r}_{13} = x_s \bar{i}_x + y_s \bar{i}_y + z_s \bar{i}_z$$

$$\bar{r}_{13} \cdot \bar{r}_{14} = \left[\frac{1}{2}(1+\rho)\, x_s + x_s x\right] + \left[\frac{\sqrt{3}}{2}(1+\rho)\, y_s + y_s y\right] + z_s z$$

Now

$$V_s = -\, m^2 \left[\frac{3}{2r_{13}^2}\left(\bar{r}_{13} \cdot \bar{r}_{14}\right)^2 - \frac{1}{2}\, r_{14}^2\right] \qquad \text{(C-9)}$$

so that only terms of $o(m^2)$ or lower must be retained inside the bracket. After dropping all terms which do not contain the particle's coordinates we get

$$\left(\bar{r}_{13} \cdot \bar{r}_{14}\right)^2 \rightarrow \left(x_s x + y_s y + \underbrace{\cancel{z_s z}}_{\text{will lead to } o(m^5) \text{ terms}}\right)^2 + \left(x_s + \sqrt{3}\, y_s\right)\left(x_s x + y_s y + z_s z\right) \qquad \text{(C-10)}$$

and

$$r_{14}^2 \rightarrow x + \sqrt{3}y + x^2 + y^2 + z^2 \qquad \text{(C-11)}$$

Substitution of (C-10) and (C-11) into (C-9) results in

$$V_s = -m^2 \left\{ \frac{3}{2r_{13}^2}\left[\left(x_s x + y_s y\right)^2 + \left(x_s + \sqrt{3}\, y_s\right)\left(x_s x + y_s y + z_s z\right)\right] - \frac{1}{2}\left[(x + \sqrt{3}y) + \left(x^2 + y^2 + z^2\right)\right] \right\} \qquad \text{(C-12)}$$

The Hamiltonian H is defined by the relation

$$H = \bar{p}_r \cdot \dot{\bar{r}} - L = p_r^T \dot{r} - L \tag{C-13}$$

where

$$\bar{p} = \left[\frac{\partial L}{\partial \dot{\bar{r}}_{14}}\right] \cdot \left[\frac{\partial \dot{\bar{r}}_{14}}{\partial \dot{\bar{r}}}\right] = \dot{\bar{r}}_{14} \cdot \underline{\underline{I}} = \dot{\bar{r}}_{14} \tag{C-14}$$

$\underline{\underline{I}}$ denotes the identity tensor. The equality $\bar{p} = \dot{\bar{r}}_{14}$ is a consequence of the linear dependence of $\dot{\bar{r}}_{14}$ on the velocity components $\dot{x},\dot{y},\dot{z}$, in the rotating coordinate frame. Writing $\dot{\bar{r}}_{14}$ as

$$\dot{\bar{r}}_{14} = \dot{\bar{r}}_{1L} + \dot{r}\bar{i}_r + \bar{\omega} \times \bar{r}$$

$$\dot{\bar{r}}_{1L} = \dot{r}_{1L}\bar{i}_{1L} + \bar{\omega} \times \bar{r}_{1L}$$

$$\dot{r}\bar{i}_r = \dot{x}\bar{i}_x + \dot{y}\bar{i}_y + \dot{z}\bar{i}_z$$

we get

$$\therefore \dot{\bar{r}}_{14} = \ldots \text{ algebra } \ldots$$

$$= \left[\frac{1}{2}\dot{\rho} - \frac{\sqrt{3}}{2}\left(1 + \upsilon_z\right)(1 + \rho) + \dot{x} + z\upsilon_y - y(1 + \upsilon_z)\right] \bar{i}_x$$

$$+ \left[\frac{\sqrt{3}}{2}\dot{\rho} + \frac{1}{2}(1 + \rho)\left(1 + \upsilon_z\right) + \dot{y} + x\left(1 + \upsilon_z\right) - z\upsilon_x\right] \bar{i}_y$$

$$+ \left[\frac{\sqrt{3}}{2}(1 + \rho)\upsilon_x - \frac{1}{2}(1 + \rho)\upsilon_y + \dot{z} + y\upsilon_x - x\upsilon_y\right] \bar{i}_z \tag{C-15}$$

We now introduce the momenta P via Eq. (19), solve for $\dot{r}$ from (C-14) and obtain the expressions

$$\dot{x} = P_x - \frac{\sqrt{3}}{2} + \left[\frac{\sqrt{3}}{2}(1+\rho) + y\right]\left(1 + \upsilon_z\right) - \frac{1}{2}\dot{\rho} - z\upsilon_y$$

$$\dot{y} = P_y + \frac{1}{2} - \frac{\sqrt{3}}{2}\dot{\rho} - \left[\frac{1}{2}(1+\rho) + x\right]\left(1 + \upsilon_z\right) + z\upsilon_x$$

$$\dot{z} = P_z + \left[\frac{1}{2}(1+\rho) + x\right]\upsilon_y - \left[\frac{\sqrt{3}}{2}(1+\rho) + y\right]\upsilon_x \qquad \text{(C-16)}$$

Also from Eq. (C-14) and Eq. (19) we can write L in the form

$$L = \frac{1}{2}\left(P_x - \frac{\sqrt{3}}{2}\right)^2 + \frac{1}{2}\left(P_y + \frac{1}{2}\right)^2 + \frac{1}{2}P_z^2 - V_{EM} - V_s \qquad \text{(C-17)}$$

If we now substitute (C-17) into (C-13), make use of (C-14) and (C-16), and neglect all the terms which do not depend on the momenta P or the particle's position $\overline{r}$ we end up after a lot of algebra with the expression for the Hamiltonian H presented in Eq. (21) of the text.

Appendix D

CANONICAL TRANSFORMATION TO SLOW VARIABLES

We shall outline here the steps which underlie the canonical transformation from the variables α,β, to the slow set α',β'. These variables are analogous to polar coordinates where $\sqrt{\alpha}$ corresponds to an amplitude and β to a phase shift. We shall find it convenient to use also a cartesian set of generalized coordinates q,p in terms of which the transformation relations will be developed.

Let

$$S = S(q,p') = S_1 + S_2 \tag{D-1}$$

be a generating function from the set q,p to a second slowly varying set q',p' where S_1 will be selected to remove from H the 3rd order terms (all of which are short period) and S_2 to remove all 4th order short period, and define

$$S_q = \left[\frac{\partial S}{\partial q_1}, \frac{\partial S}{\partial q_2}, \frac{\partial S}{\partial q_3}\right] = (1 \times 3) \text{ row matrix of partial derivatives of S}$$

S_{q^T} = (3 × 1) column matrix of partial derivatives

Then

$$\begin{aligned} p &= p' + S_{q^T}(q,p') \\ q &= q' - S_{p'^T}(q,p') \end{aligned} \tag{D-2}$$

Let

$$\begin{aligned} \Delta q &= q - q' \\ \Delta p &= p - p' \end{aligned} \tag{D-3}$$

and expand S_{q^T} and $S_{p'^T}$ in a Taylor series around the values of q', p'. To second order in Δq and Δp

$$S_{q^T}(q,p') = \frac{\partial}{\partial q^T}\left[S(q',p') + S_{q'}\Delta q + \frac{1}{2}\,\Delta q^T S_{q'^T q'}\,\Delta q + \cdots\right] \tag{D-4}$$

The expressions for $q = q(q',p')$ and $p = p(q',p')$ can then be developed via (D-2) and (D-4)

$$q = q' - S_{p'^T}(q,p') = q' - \frac{\partial}{\partial_{p'^T}}\left[S(q',p') + S_{q'}\Delta q + \frac{1}{2}\,\Delta q^T S_{q'^T q'}\,\Delta q + \cdots\right]$$

$$\simeq q' - S_{p'^T}(q',p') - S(q',p')_{p'^T q'}\Delta q + \cdots \tag{D-5}$$

In order to prevent carrying unnecessary terms along let us estimate the order of magnitude of the above terms. Since S will be used to perform a transformation of variables in H' which contains terms of $o(m^3)$ and $o(m^4)$ then the lowest terms in S will be of $o(m^3)$. We might also use the notation $o(x^3)$ since the x,y,z coordinates are the ones to be transformed.

Let us view q as equivalent to β and p as equivalent to α.

Then the derivatives result in the following orders of magnitude

$$\begin{aligned} S &= o(H_3) \simeq o(x^3) = o(\alpha^{3/2}) \\ S_p &\to S_\alpha = o(\alpha^{1/2}) = o(x) \\ S_q &\to S_\beta = o(\alpha^{3/2}) = o(x^3) \\ S_{p^T q} &\to o(S_p) = o(x) \end{aligned} \tag{D-6}$$

We assume also that

$$\Delta q \cong o(x) + \text{higher order terms} \tag{D-7}$$

We thus note that the term $\Delta q^T S_{q^T q'}\, \Delta q = o(x^5)$ and after operating on it with ∂/∂_p, it becomes $o(x^3)$.

Equation (D-5) can then be written as

$$\Delta q = - S_{p^T} - S_{p^T q'}\, \Delta q + o(x^3) \cdots \tag{D-8}$$

and terms of $o(x^3)$ are not carried along. Thus to $o(x^2)$ we can write the following relation

$$\left[I + S_{p^T q'}\right] \Delta q \cong - S_{p^T}$$

which can be inverted to solve for Δq

$$\Delta q \cong - \left[I + S_{p^T q'}\right]^{-1} S_{p^T} \cong - S_{p^T} + S_{p^T q'}\, S_{p^T} + o(x^3) \tag{D-9}$$

where I is the identity matrix.

From Eq. (D-2) we also note that

$$\Delta p \approx o(S_{q^T}) = o(x^3) + \text{H.O.T.}$$

Expanding for Δp as was done in (D-9) for Δq we find

$$\Delta p \cong S_{q^T} - S_{q^T q'}\, S_{p^T} + o(x^5) \tag{D-10}$$

The partial derivatives of H can also be treated similarly to the partials of S. Thus

$$H_q = o(H)$$
$$H_p = o\left(\frac{H}{x^2}\right) \qquad \text{(D-11)}$$

For a scleronomic generating function S we have the transformation relation for Hamiltonians

$$K(q',p') = H(q,p) \qquad \text{(D-12)}$$

and expanding H in a Taylor series around q',p' gives

$$\begin{aligned}
K(q',p') &= H(q',p') + H_{q'}\,\Delta q + H_{p'}\,\Delta p + \\
&\quad + \frac{1}{2!}\left[\Delta q^T \frac{\partial}{\partial q'^T} + \Delta p^T \frac{\partial}{\partial p'^T}\right]^2 H(q',p') + \cdots \\
&= H(q',p') + H_{q'}\Delta q + H_{p'}\Delta p + \frac{1}{2}\left[\Delta q^T H_{q'^T q'}\,\Delta q\right. \\
&\quad \left. + \Delta q^T H_{q'^T p'}\,\Delta p + \Delta p^T H_{p'^T q'}\,\Delta q + \Delta p^T H_{p'^T p'}\,\Delta p\right] + \cdots \\
&= H(q',p') + H_{q'}\left[- S_{p'^T} + S_{p'^T q'} S_{p'^T}\right]_A + H_{p'}\left[S_{q'^T} - S_{q'^T q'} S_{p'^T}\right]_B \\
&\quad + \frac{1}{2}\left[- S_{p'} + S_{p'} S_{p'q'^T}\right]_C H_{q'^T q'}\left[- S_{p'^T} + S_{p'^T q'} S_{p'^T}\right] \\
&\quad + \frac{1}{2}\left[- S_{p'} + S_{p'} S_{p'q'^T}\right]_D H_{q'^T p'}\left[S_{q'^T} - S_{q'^T q'} S_{p'^T}\right] \\
&\quad + \frac{1}{2}\left[S_{q'} - S_{p'} S_{q'q'^T}\right] H_{p'^T q'}\left[- S_{p'^T} + S_{p'^T q'} S_{p'^T}\right] \\
&\quad + \frac{1}{2}\left[S_{q'} - S_{p'} S_{q'q'^T}\right] H_{p'^T p'}\left[S_{q'^T} - S_{q'^T q'} S_{p'^T}\right] + \cdots
\end{aligned}$$
$$\text{(D-13)}$$

The subscript letters A,B,C,D have been introduced merely for ease of subsequent identification of the respective brackets $[\ ]$.

We recall that

$$H = \underset{o(x^2)}{H^{(o)}} + \underset{o(x^3)}{H_3} + \underset{o(x^4)}{H_4} \tag{D-14}$$

where H_3 contains x^3 terms and H_4 denotes the 4th order terms like x^4, m^2x^2, ρx^2 etc. This breakdown into H_3 and H_4 will be made use of in choosing the relations defining $S_1(q',p')$ and $S_2(q',p')$.

We shall assume S_1 to contain only $o(x^3)$ terms and S_2 only terms of $o(x^4)$, because we shall select S_2 so as to remove all 4th order short period terms from the Hamiltonian K.

$$[\ ]_A \rightarrow -\left[S_{1_{p'T}} + S_{2_{p'T}} - \left(S_{1_{p'T_{q'}}} + S_{2_{p'T_{q'}}}\right)\left(S_{1_{p'T}} + S_{2_{p'T}}\right)\right]$$

$$= -\underset{o(x)}{S_{1_{p'T}}} - \underset{o(x^2)}{S_{2_{p'T}}} + \underset{o(x^2)}{S_{1_{p'T_{q'}}}S_{1_{p'T}}} + \underset{o(x^3)}{S_{1_{p'T_{q'}}}S_{2_{p'T}}} + \underset{o(x^3)}{S_{2_{p'T_{q'}}}S_{1_{p'T}}} + o(x^4)$$

(D-15)

Thus to $o(x^4)$, which is the highest order retained in all terms,

$$H_{q'}[\ ]_A \rightarrow H^{(o)}_{q'}\left[-S_{1_{p'T}} - S_{2_{p'T}} + S_{1_{p'T_{q'}}}S_{1_{p'T}}\right] + H_{3_{q'}}\left[-S_{1_{p'T}}\right] \tag{D-16}$$

Similarly, the following expressions can be derived

$$[\ ]_B \rightarrow S_{1_{q'T}} + S_{2_{q'T}} - S_{1_{q'T_{q'}}}S_{1_{p'T}} + \text{H.O.T.} \tag{D-17}$$

$$H_{p},[\ ]_B \rightarrow H^{(o)}_{p},\left[S_{1_{q,T}} + S_{2_{q,T}} - S_{1_{q,T_q}}, S_{1_{p,T}}\right] + H_{3_{p}}, S_{1_{q,T}} + \text{H.O.T.} \tag{D-18}$$

$$\tfrac{1}{2}[\ ]_C\, H_{q,T_q}, [\ ]_A \rightarrow \tfrac{1}{2} S_{1_{p}}, H^{(o)}_{q,T_q}, S_{1_{p,T}} \tag{D-19}$$

$$\tfrac{1}{2}[\ ]_C\, H_{q,T_p}, [\ ]_B \rightarrow -\tfrac{1}{2} S_{1_{p}}, H^{(o)}_{q,T_p}, S_{1_{q,T}} \tag{D-20}$$

$$\tfrac{1}{2}[\ ]_D\, H_{p,T_q}, [\ ]_A \rightarrow -\tfrac{1}{2} S_{1_{q}}, H^{(o)}_{p,T_q}, S_{1_{p,T}} \tag{D-21}$$

$$\tfrac{1}{2}[\ ]_D\, H_{p,T_p}, [\ ]_B \rightarrow \tfrac{1}{2} S_{1_{q}}, H^{(o)}_{p,T_p}, S_{1_{q,T}} \tag{D-22}$$

Substituting (D-16) and (D-18) through (D-22) into (D-13) results in the expression

$$\begin{aligned} K = {} & H^{(o)} + H_3 + H_4 + H^{(o)}_{q},\left[- S_{1_{p,T}} - S_{2_{p,T}} + S_{1_{p,T_q}}, S_{1_{p,T}}\right] \\ & - H_{3_{q}}, S_{1_{p,T}} + H^{(o)}_{p},\left[S_{1_{q,T}} + S_{2_{q,T}} - S_{1_{q,T_q}}, S_{1_{p,T}}\right] + H_{3_{p}}, S_{1_{q,T}} \\ & + \tfrac{1}{2} S_{1_{p}}, H^{(o)}_{q,T_q}, S_{1_{p,T}} - \tfrac{1}{2} S_{1_{p}}, H^{(o)}_{q,T_p}, S_{1_{q,T}} - \tfrac{1}{2} S_{1_{q}}, H^{(o)}_{p,T_q}, S_{1_{p,T}} \\ & + \tfrac{1}{2} S_{1_{q}}, H^{(o)}_{p,T_p}, S_{1_{q,T}} \end{aligned} \tag{D-23}$$

Recalling the definition of the Poisson bracket

$$[H,S] = H_q S_p{}^T - H_p S_q{}^T \tag{D-24}$$

and applying it to the terms of Eq. (D-23) one can obtain after a lengthy series of manipulations and combinations of terms the expression

$$\begin{aligned} K = H^{(o)} + H_3 + H_4 - \left[H^{(o)},S_1\right] - \left[H^{(o)},S_2\right] - \left[H_3,S_1\right] \\ + \frac{1}{2}\left[H^{(o)},S_{1_q}{}', S_{1_p}{}'^T\right] + \frac{1}{2}\left[\left[H^{(o)},S_1\right], S_1\right] \end{aligned} \tag{D-25}$$

in terms of q' and p' only.

Let us choose for the definition of S_1 the relation

$$\left[H^{(o)},S_1\right] - H_3 = 0 \tag{D-26}$$

and thereby remove H_3 which contains only short period terms.

Then

$$\begin{aligned} K = H^{(o)} + H_4 + \frac{1}{2}\left[H^{(o)},S_{1_q}{}', S_{1_p}{}'^T\right] + \frac{1}{2}\left[H_3,S_1\right] - \left[H_3,S_1\right] \\ - \left[H^{(o)},S_2\right] = H^{(o)} + H_4 + \frac{1}{2}\left[H^{(o)},S_{1_q}{}', S_{1_p}{}'^T\right] \\ - \frac{1}{2}\left[H_3,S_1\right] - \left[H^{(o)},S_2\right] \end{aligned} \tag{D-27}$$

The third and fifth final terms in Eq. (D-27) can be combined into the one bracket

$$- \left[H^{(o)}, S_2 - \frac{1}{2} S_{1_{q}}, S_{1_{p}} \right]_T \qquad \text{(D-28)}$$

We now define S_2 such as to remove all remaining 4th order short period terms from K. Now both H_4 and $[H_3,S_1]$ will contain both long period $\overline{(\ )}$ and short period (s·p) terms, which can all be eliminated by letting S_2 be defined via

$$H_{4_{s \cdot p}} - \frac{1}{2} \left[H_3, S_1 \right]_{s \cdot p} - \left[H^{(o)}, S_2 - \frac{1}{2} S_{1_{q}}, S_{1_{p}} \right]_T = 0 \qquad \text{(D-29)}$$

This leaves the long period form of the Hamiltonian K as

$$K = H^{(o)} + \overline{H}_4 - \frac{1}{2} \overline{\left[H_3, S_1 \right]} \qquad \text{(D-30)}$$

and if α' and β' are selected as the canonical variables, rather than α' and $(t + \beta')$, the long period perturbation Hamiltonian K' is obtained as

$$K' = \overline{H}_4 - \frac{1}{2} \overline{\left[H_3, S_1 \right]} \qquad \text{(D-31)}$$

To this expression one must still add the contribution from the linear forced solutions $\tilde{x},\tilde{y}$ due to $H^{(o)}$ as indicated in Eq. (49) which then finally leads to the relation presented in Eq. (48).

Comparison of the K from Eq. (D-31) with the K presented on p. 63 of Ref. 6 shows that the two Hamiltonians are not alike. This difference can be traced to the particular way in which the time dependent generating function $S_1(q,p',t)$ of Ref. 6 was defined there by means of an equation in the mixed variables q,p', instead of first carrying out the transformation to the new set of coordinates q',p' shown in this appendix in Eqs. (D-9) to (D-13).

As a consequence of the use of mixed variables, some of the terms which would have appeared from the additional Taylor series expansion over q were thus missing and only one half of the terms of the Poisson bracket $[H_3, S_1]$ of Eq. (D-31) showed up in the function Φ introduced in Ref. 6. The absence of these additional terms prevented the cancellation of nonpolynomial terms (i.e., terms which do not arise from binomial expansions such as $(x + y)^n$, where n is some finite integer) and led to the presence of an extraneous term such as the $\alpha_1'^{3/2}\alpha_2'^{1/2}$ term in Eq. (10) of Ref. 6.

The source of the incorrect results, which can arise when one operates with mixed variables anytime terms higher than of first order are retained in the Hamiltonian, were recognized by Prof. Breakwell, who then suggested that the correct procedure in choosing the function S_1 would be to transform first to the new set of coordinates q', p'. The implementation of this suggestion led to the developments presented in this appendix, and avoided here the presence of the inadmissible nonpolynomial terms.

The derivation of the Mathieu type Hamiltonian in Appendix F does make use of mixed variables. However, the results obtained there are correct since only linear terms were retained in H.

A last comment should be made regarding the slow variables q', p', or α', β'. It turns out that it is impossible to prevent the presence of some higher order long period terms in S_2 which arise because the term $S_{1_{q'}} S_{1_{p'T}}$ may contain also long period parts. From this it follows that in the expression for, say, q

$$q = q' - S_{1_{p'T}} + S_{1_{p'Tq'}} S_{1_{p'T}} - S_{2_{p_1'T}} \tag{D-32}$$

the last two terms may also make some long period contributions to q, which would tend to contradict the assertion that q' (and also p') are the only long period variables. This situation is unfortunately unavoidable and cannot be circumvented by redefining S_1 or S_2, since the elimination of the extra long period terms in q' or p' via S would automatically result in the introduction of unwanted higher order short period terms into K that S would be incapable of suppressing simultaneously.

Fortunately this impasse is not too serious since the bothersome long period terms in Eq. (D-32) are of $o(m^4)$ or higher and may be safely disregarded within the extent of the present theory inasmuch as q' does not appear in a linear manner in H. They would pose a problem however if the present approach were to be extended to encompass some of the higher order terms currently neglected.

Appendix E

SOME ILLUSTRATIVE STEPS IN THE DERIVATION OF LONG PERIOD TERMS IN K′

The steps leading from Eq. (48) to Eq. (50) required by far the most time consuming, tedious and exacting manipulations and computations of the whole investigation. We shall indicate here only briefly as an example a few representative intermediate steps so as to provide the reader with a feeling for what is involved here.

First a general remark concerning the Poisson bracket $\overline{[H_3,S_1]}$. In the expanded form, and using the polar canonical variables α and β, this becomes

$$\left[H_3,S_1\right] = H_{3\beta_i'}S_{1\alpha_i'} - H_{3\alpha_i'}S_{1\beta_i'} \qquad \text{(E-1)}$$

where the tensor notation for summation over i = 1,2,3 has been used. The same bracket, when H_3 and S_1 are expressed in cartesian coordinates x,y,z,P_x,P_y,P_z, can also be written as

$$\left[H_3,S_1\right] = H_{3x}S_{1P_x} + H_{3y}S_{1P_y} + \cdots - H_{3P_x}S_{1x} - \cdots - H_{3P_z}S_{1z} \qquad \text{(E-2)}$$

which indicates that the bracket will give rise only to polynomial terms of the form $x^2P_z^2$, x^3P_y, y^4, etc.

From this it follows that when one evaluates the long period terms in the polar coordinates used in Eq. (E-1) one must be careful to observe that only polynomial type terms should be retained. Thus, one can obtain secular terms like $5\alpha_1'^2$, $7\alpha_3'^2$ ··· etc. or slowly varying terms like $(\cdots)\ \alpha_1'\alpha_3' \cos[(\omega_1 - \omega_3)t + \cdots]$ or $(\cdots)\ \alpha_1'^{1/2}\alpha_2'^{3/2} \times \cos[(\omega_1 - 3\omega_2)t + \cdots]$, but not terms such as $(\cdots)\ \alpha_1'^{3/2}\alpha_2'^{1/2} \times \cos[(\omega_1 - 3\omega_2)t + \cdots]$ because such a term could not arise from products of the form $x_1y_2^3$ or $y_1x_2^3$ which are the only kind that could give rise to long period trigonometric terms with a frequency $\omega_1 - 3\omega_2$. The quantities x_1,y_2, etc. represent the ω_1 term in $\bar{x}$ and the ω_2 term in $\bar{y}$ of Eq. (36), respectively.

This polynomial requirement is not satisfied in Eq. (10) of Ref. (6) which contains the term $16.2\alpha_1'^{3/2}\alpha_2'^{1/2}\cos[.0609t + \omega_1\beta_1' + 3\omega_2\beta_2' - 4.48^{\circ}]$.

We shall indicate now a few steps in the evaluation of one of the long period terms in the Poisson bracket. For convenience we let

$$H_3 = H_{32} + H_{33} \tag{E-3}$$

where

H_{32} = coplanar (x,y) terms in H_3
H_{33} = out-of-plane (z) terms in H_3

Similarly

$$S_1 = S_{12} + S_{13} \tag{E-4}$$

where

$$S_1 = -\int^t H_3 dt \tag{E-5}$$

from Appendix D. (We recall that $[H,S] = -\partial S/\partial t$ when H is treated as the momentum conjugate to the coordinate t.)

Then

$$\overline{[H_3, S_1]} = \overline{[H_{32} + H_{33}, S_{12} + S_{13}]} =$$

$$= \overline{[H_{32}, S_{12}]} + \overline{[H_{32}, S_{13}]} + \overline{[H_{33}, S_{12}]} + \overline{[H_{33}, S_{13}]} \tag{E-6}$$

Let us take the first bracket in Eq. (E-6), and consider for example only the component $H_{\alpha_1}S_{1\beta_1}$ in it. It can be shown that it arises from the product of the two parts

$$H_{32\alpha_1} = \alpha_1^{1/2}\left\{\frac{3}{2} M(a_1 + a_2) + \frac{3}{2} N(33a_3 - 7a_4)\right\}_A$$

$$+ \alpha_2^{1/2}\left\{M(b_1 + b_2) + N(33b_3 - 7b_4)\right\}_B$$

$$+ \alpha_1^{-1/2}\alpha^2\left\{\frac{1}{2} M(c_1 + c_2) + \frac{1}{2} N(33c_3 - 7c_4)\right\}_C \qquad \text{(E-6)}$$

and

$$S_{12\beta_1} = -\alpha_1^{3/2}\left\{M(a_1' + a_2') + N(33\alpha_3' - 7\alpha_4')\right\}_D$$

$$- \alpha_1\alpha_2^{1/2}\left\{M(b_1' + b_2') + N(33b_3' - 7b_4')\right\}_E$$

$$- \alpha_1^{1/2}\alpha_2\left\{M(c_1' + c_2') + N(33c_3' - 7c_4')\right\}_F$$

where we have dropped for convenience the primes on the α's and β's.

The quantities $a_1 \cdots a_4$, $b_1 \cdots b_4$, $\cdots c_1' \cdots c_4'$ are defined in terms of $\overline{x} = A_1C_{(1)} + A_2C_{(2)}$ and $\overline{y} = A_1'C_{(1)+\delta_1} + A_2'C_{(2)+\delta_2}$, where $(1) = \omega_1\beta_1^{\neq}$, $(2) = \omega_2\beta_2^{\neq}$ and $A_1, A_1' \cdots \delta_1, \delta_2$ are obtained from Eq. (36). M and N are two constants defined as $M = 3\sqrt{3}/16$ and $N = (1 - 2\mu)/16$.

In terms of the above constants one can obtain the following expressions

$$a_1 = a_1' = \frac{1}{2} A_1^2 A_1^1\left(C_{(1)+\delta_1} + \frac{1}{2} C_{3(1)+\delta_1} + \frac{1}{2} C_{(1)-\delta_1}\right)$$

$$b_1' = A_1A_1'A_2\left[\frac{\omega_1}{2\omega_1 + \omega_2} C_{2(1)+(2)+\delta_1} + \frac{\omega_1}{2\omega_1 - \omega_2} C_{2(1)-(2)+\delta_1}\right]$$

$$+ \frac{1}{2} A_1^2 A_2^1\left[\frac{\omega_1}{2\omega_1 + \omega_2} C_{2(1)+(2)+\delta_2} + \frac{\omega_1}{2\omega_1 - \omega_2} C_{2(1)-(2)-\delta_2}\right] \qquad \text{(E-7)}$$

$\vdots$

and similar expressions for all the other quantities inside the ( ) brackets.

We observe from Eq. (E-6) that in the Poisson bracket, the coefficient of α_1^2 would arise from the product of brackets $\{\ \}_A$ with $\{\ \}_D$, and in the same manner we note that the

$$
\begin{array}{lll}
\text{coefficient of} & \alpha_1\alpha_2 \rightarrow \text{results from} & \{\ \}_A\cdot\{\ \}_F;\ \{\ \}_B\cdot\{\ \}_E;\ \{\ \}_C\cdot\{\ \}_D \\
\quad '' & \alpha_1^{1/2}\alpha_2^{3/2} \rightarrow \text{results from} & \{\ \}_B\cdot\{\ \}_F;\ \{\ \}_C\cdot\{\ \}_E \\
\quad '' & \alpha_2^2 \rightarrow \text{results from} & \{\ \}_C\cdot\{\ \}_F
\end{array}
\qquad \text{(E-8)}
$$

Products of brackets $\{\ \}$ as indicated in Eq. (E-8) arise in all the partial derivative products of H_3 with S_1, and must be summed up for every combination $\alpha_i^n\alpha_j^m$ to obtain the final value of the coefficient for that particular combination of α's.

For example, to obtain the coefficient of α_1^2 in $H_{32\alpha_1}S_{12\beta_1}$ we have

$$
\alpha_1^2:\quad -\{\ \}_A\{\ \}_D = -\frac{3}{2}\left\{M(a_1 + a_2) + N(33a_3 - 7a_4)\right\}
$$

$$
\cdot\left\{M(a_1' + a_2') + N(33a_3' - 7a_4')\right\} \qquad \text{(E-9)}
$$

where the following relations among the a's apply here

$$
\begin{array}{ll}
a_1 = a_1' & a_3 = a_3' \\
a_2 = a_2' & a_4 = a_4'
\end{array}
\qquad \text{(E-10)}
$$

Expanding and using the appropriate relations for the a's (not shown here) gives

$$\alpha_1^2 :\ - \left\{\frac{3}{2} M^2 (a_1 + a_2)^2 + 3MN(a_1 + a_2)(33a_3 - 7a_4) + \frac{3}{2} N^2 (33a_3 - 7a_4)^2 \right\} \tag{E-11}$$

$$\begin{aligned} \frac{3}{2} M^2 (a_1 + a_2)^2 &= \frac{3}{2} M^2 (a_1^2 + 2a_1 a_2 + a_2^2) \\ &= \frac{3}{2} M^2 \left[\frac{1}{4} A_1^4 A_1'^2 \left(\frac{3}{4} + \frac{1}{2} C_{2\delta_1}\right) + \frac{1}{4} A_1^2 A_1'^4 \left(\frac{3}{2} + C_{2\delta_1}\right) + \frac{1}{16} A_1'^6 \cdot 5\right] \end{aligned} \tag{E-12}$$

$$\begin{aligned} 3MN(a_1 + a_2)(33a_3 - 7a_4) = 3MN &\left[33\left\{\frac{1}{4} A_1^3 A_1'^3 \left(\frac{9}{8} C_{\delta_1} + \frac{1}{8} C_{3\delta_1}\right) + \frac{1}{8} A_1 A_1'^5 \cdot \frac{10}{4} C_{\delta_1}\right\}\right. \\ &\left. - 7\left\{\frac{1}{8} A_1^5 A_1' \frac{10}{4} C_{\delta_1} + \frac{1}{16} A_1^3 A_1'^3 \left(\frac{9}{2} C_{\delta_1} + \frac{1}{2} C_{3\delta_1}\right)\right\}\right] \end{aligned} \tag{E-13}$$

$$\begin{aligned} \frac{3}{2} N^2 (33a_3 - 7a_4)^2 = \frac{3}{2} N^2 &\left[1089\ \frac{1}{4}\left\{A_1^2 A_1'^4 \left(\frac{3}{4} + \frac{1}{2} C_{2\delta_1}\right)\right\} - 462\left\{\frac{1}{8} A_1^4 A_1'^2 \left(\frac{3}{2} + C_{2\delta_1}\right)\right\}\right. \\ &\left. + 49\left\{\frac{1}{16} A_1^6 \cdot 5\right\}\right] \end{aligned} \tag{E-14}$$

When the numerical values for $A_1, A_1', A_2, A_2', C_{\delta_1}, C_{2\delta_1}$, etc. are substituted into Eqs. (E-12) through (E-14) and all the terms added, one obtains the result

$$-92.871\alpha_1^2 \tag{E-15}$$

This same, or a similar, procedure must be repeated for every combination of α's which arises from all the terms of the Poisson bracket.

Appendix F

MATHIEU TYPE HAMILTONIANS

Consider the near-resonant Mathieu equation

$$\ddot{X} + \omega_o^2\left[1 + \eta \cos 2\omega_o t \; (1 + \epsilon)\right]X = 0 \tag{F-1}$$

where η, ϵ, $<< 1$. The effect of the trigonometric coefficient is to introduce a parametric excitation term into the simple harmonic oscillator model.

It is easily seen that as $\epsilon \to 0$ a resonant forcing term will arise in case a perturbation solution is attempted, after the $X^{(o)}$ solution to the equation

$$\ddot{X} + \omega_o^2 X = 0 \tag{F-2}$$

is substituted back into Eq. (F-1) to provide the next higher term.

The Hamiltonian of system (F-1) is

$$\begin{aligned} H &= \frac{1}{2}\left[p^2 + \omega^2 X^2\right] = H^{(o)} + H' \\ \omega &= \omega_o\left[1 + \eta \cos 2\omega_o t(1 + \epsilon)\right]^{1/2} \end{aligned} \tag{F-3}$$

X, p = generalized coordinate and momentum, respectively

$H^{(o)}$ = Hamiltonian of simple harmonic oscillator of frequency ω_o

H' = perturbation Hamiltonian (for $\eta << 1$)

The solution corresponding only to $H^{(o)}$ is

$$X^{(o)} = \frac{\sqrt{2\alpha}}{\omega_o} \sin \omega_o \; (t + \beta) \tag{F-4}$$

where α, β are constants of integration.

When H' is included, α and β become functions of time t. It is useful to consider mainly the long period variations in α and β, since these basically tell us the most about the "averaged" long term behavior of the system.

The canonical transformations of variables shown next serve the purpose of suppressing all short period terms in H. We assume that α and β can be decomposed into short period and long period (α', β') components, i.e.,

$$\alpha \rightarrow \alpha' + \alpha_{s.p.}$$

$$\beta \rightarrow \beta' + \beta_{s.p.}$$

Introducing $X^{(o)}$ into the Hamiltonian H'

$$H' = \frac{1}{2} \omega_o^2 \eta X^2 \cos 2\omega_o t \, (1 + \epsilon) \qquad \text{(F-5)}$$

and rearranging terms gives

$$H' = \frac{\alpha\eta}{2} \left\{ \cos 2\omega_o t(1 + \epsilon) - \frac{1}{2} \cos \left[2\omega_o t(1 + \epsilon) + 2\omega_o(t + \beta) \right] - \frac{1}{2} \cos 2\omega_o(\epsilon t - \beta) \right\} \qquad \text{(F-6)}$$

The last term with angular velocity $2\omega_o\epsilon << 2\omega_o$ is of low frequency and thus gives a contribution to the long period part of H'.

Let this term be designated by $\overline{H}'$

$$\overline{H}' = - \frac{\alpha\eta}{4} \cos 2\omega_o(\epsilon t - \beta) \qquad \text{(F-7)}$$

Note that α still contains s.p. terms so that $\overline{H}'$ still is not the final form of the desired long period Hamiltonian.

To obtain the D.E. for α' we introduce a generating function $S(\alpha',\beta,t)$ of the Hamilton-Jacobi equation which, to first order terms can be written in the form

$$S = \alpha'\beta + \eta S_1(\alpha',\beta,t) \tag{F-8}$$

in terms of the new momenta α' and the old coordinates β. Thus

$$\begin{aligned} \alpha &= \frac{\partial S}{\partial \beta} = \alpha' + \eta\,\frac{\partial S_1}{\partial \beta} \\ \beta' &= \frac{\partial S}{\partial \alpha'} = \beta + \eta\,\frac{\partial S_1}{\partial \alpha'} \end{aligned} \tag{F-9}$$

relate the old and new coordinates and momenta. α',β' form a canonical set with respect to a long period Hamiltonian K', such that

$$\begin{aligned} \dot{\alpha}' &= -\frac{\partial K'}{\partial \beta'} \\ \dot{\beta}' &= \frac{\partial K'}{\partial \alpha'} \end{aligned} \tag{F-10}$$

where

$$K' = H' + \frac{\partial S}{\partial t} \tag{F-11}$$

The Hamiltonian K' of Eq. (F-11) is treated as a function of the coordinates α',β' and t, after the transformation relations (F-9) are substituted into the right hand side of (F-11).

To linear terms only

$$H'(\alpha,\beta,t) = H'(\alpha' + \eta S_{1\beta},\beta,t) \cong H'(\alpha',\beta,t) + \text{H.O.T.} \tag{F-12}$$

and thus

$$K' \simeq H'_{sp}(\alpha', \beta, t) + \overline{H}'(\alpha', \beta, t) + \eta S_{1_t} + \text{H.O.T.} \tag{F-13}$$

The function S_1 is chosen in such a way as to eliminate all s.p. terms from (F-13). We thus require

$$\eta S_{1_t} + H'_{sp}(\alpha', \beta, t) = 0 \tag{F-14}$$

from which results

$$S_1 = \frac{\alpha'}{8\omega_o(2+\epsilon)} \sin 2\omega_o(2t + \epsilon t + \beta) - \frac{\alpha'}{4\omega_o(1+\epsilon)} \sin 2\omega_o t(1+\epsilon) \tag{F-15}$$

Expanding β around β' in $\overline{H}'$ of Eq. (F-13) gives, again to first order

$$K' = \overline{H}'(\alpha', \beta', t) = -\frac{\alpha'\eta}{4} \cos 2\omega_o(\epsilon t - \beta') \tag{F-16}$$

with the aid of Eq. (F-7).

Equation (F-16) defines a long period, time dependent, Mathieu type Hamiltonian.

<u>Stability Analysis</u>

The differential equations for α' and β' are summarized by the matrix equation

$$\begin{pmatrix} \dot{\beta}' \\ \dot{\alpha}' \end{pmatrix} = \Phi_o \begin{pmatrix} \partial/\partial\beta' \\ \partial/\partial\alpha' \end{pmatrix} K' \tag{F-17}$$

Rather than solve Eq. (F-17) directly for α' and β' it is more convenient to introduce a further generating functions S^* (α^*, β', t)

so as to eliminate the explicit dependence on time and thereby transform K' to a new Hamiltonian K^* which is a constant of the motion, i.e., $K^* = K_o^* =$ constant.

Let us define a variable β^* by the relation

$$\beta^* = \beta' - \epsilon t \tag{F-18}$$

and then take S^* to be given by

$$S^* = \alpha^*(\beta' - \epsilon t) \tag{F-19}$$

Since

$$S^*_{\alpha^*} = \beta' - \epsilon t = \beta^*$$

and

$$S_{\beta'} = \alpha' = \alpha^*$$

the two variables α^* and β^* are canonically related to the new Hamiltonian K^* which becomes

$$K^* = K' + S^*_t = K' - \epsilon\alpha^* = -\frac{\alpha^*\eta}{4}\cos 2\omega_o\beta^* - \epsilon\alpha^* \tag{F-20}$$

That K^* is an integral constant of the motion is evident from the fact that

$$\frac{dK^*}{dt} = \dot{K}^* = \frac{\partial K^*}{\partial\alpha^*}\dot{\alpha}^* + \frac{\partial K^*}{\partial\beta^*}\dot{\beta}^* + \cancel{\frac{\partial K^*}{\partial t}}$$

$$= \text{by Hamilton's equation} = \frac{\partial K^*}{\partial\alpha^*}\left(-\frac{\partial K^*}{\partial\beta^*}\right) + \frac{\partial K^*}{\partial\beta^*}\left(\frac{\partial K^*}{\partial\alpha^*}\right) = 0$$

$$\therefore K^* = K_o^* = \text{constant of the motion}$$

The stability or instability of systems governed by Hamiltonians of the form (F-20) can readily be established based on a comparison of the magnitude of the coefficients of α^* and $\alpha^* \cos 2\omega_o\beta^*$ in (F-20).

The relative magnitude required of these coefficients for instability or stability to exist can be determined as shown below, and the conclusions then checked by referring to the known stability regions of the Mathieu plane.

From Eq. (F-20) we obtain the differential equations for α^*

$$\dot{\alpha}^* = -\frac{\partial K^*}{\partial\beta^*} = -\frac{\omega_o\alpha^*\eta}{2}\sin 2\omega_o\beta^* \qquad \text{(F-21)}$$

and squaring,

$$\dot{\alpha}^{*2} = \frac{\omega_o^2\alpha^{*2}\eta^2}{4}\left\{1 - \left[\frac{4}{\alpha^*\eta}K_o^* + \frac{4\epsilon}{\alpha^*\eta}\alpha^*\right]^2\right\} \qquad \text{(F-22)}$$

after $\sin^2 2\omega_o\beta^*$ is replaced from Eq. (F-20) and the constancy of K^* is made use of.

The condition necessary for $\dot{\alpha}^*$ to vanish is obtained by setting the right hand side of (F-22) equal to zero, i.e., at the intersection of the two lines.

$$\left.\begin{aligned} &y = \pm\,\alpha^* \\ \text{and}\quad & \\ &y = \frac{4}{\eta}K_o^* + \frac{4\epsilon}{\eta}\alpha^* \end{aligned}\right\} \qquad \text{(F-23)}$$

This is shown in the next sketch, Fig. (F-1).

From this sketch we see that for $\alpha_o^* > \alpha_{cr}^*$ and $\dot{\alpha}_o^* > 0$ the variation of α^* is bounded by the lines $y = \pm\,\alpha^*$ if $4\epsilon/\eta > 1$, thus implying a stable motion, while if $4\epsilon/\eta < 1$, α^* grows without limit.

Hence, if

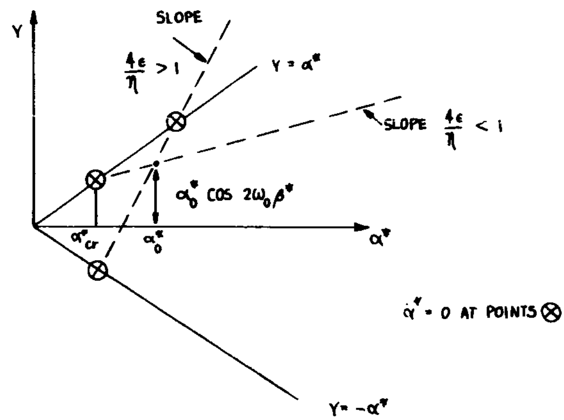


Fig. (F-1): Stability Conditions for Mathieu Type Hamiltonians

$\frac{4\epsilon}{\eta} > 1 \rightarrow$ stability exists
(i.e., x = bounded)

$\frac{4\epsilon}{\eta} < 1 \rightarrow$ instability exists
(i.e., $x \rightarrow \infty$ as $t \rightarrow \infty$)

This leads to the conclusion that the motion is unstable if in the Hamiltonian K^* the coefficient of α^* is smaller than the coefficient of $\alpha^* \cos 2\omega_o\beta^*$.

The above conclusion is also borne out by considering the Mathieu Equation in the standard form,

$$\frac{d^2v}{dz^2} + (a - 2q \cos 2z)v = 0 \qquad \text{(F-24)}$$

Referring to Eq. (F-1) and introducing a new time variable τ

$$\tau = \omega_o t(1 + \epsilon) + \frac{\pi}{2}$$

and (F-25)

$$\frac{d}{dt} = \omega_o (1 + \epsilon) \frac{d}{d\tau}$$

Equation (F-1) reduces to (F-24) if

$$a = \frac{1}{(1 + \epsilon)^2}$$

and (F-26)

$$q = \frac{\eta}{2(1 + \epsilon)^2} = \frac{\eta}{2} a$$

The stability boundaries of the Mathieu plane (q,a) in the vicinity of the region $a \simeq 1$ are shown below (see for instance p. 114 of Ref. 11).

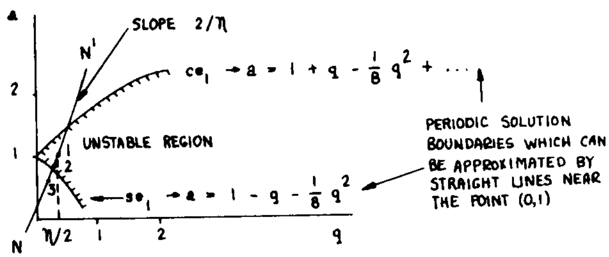


Fig. (F-2): Stability Boundaries in Mathieu Plane (q,a)

On se_1 the slope $(da/dq)_{q\to 0} \rightarrow -1$.

The slope of line N - N′ is, from Eq. (F-26), $(da/dq)_{N-N'} = 2/\eta$, and $q = \eta/2$ for $a = 1$ (pt. 1).

As the value of (a) changes from 1 to $1/(1+\epsilon)^2 \simeq 1 - 2\epsilon + 3\epsilon^2$, point 1 translates along line $N - N'$ to either point 2 or 3, depending on whether the solution remains unstable (pt. 2) or enters the stable region (pt. 3).

Comparing the values of q on se_1 and $N\text{-}N'$ for the same change $\Delta a \simeq -2\epsilon$ + H.O.T. we have

$$\underline{\text{On } se_1} \qquad q_{se} \simeq \left(\frac{dq}{da}\right)_{q=0} \Delta a = (-1)(-2\epsilon) = 2\epsilon$$

$$\underline{\text{On } N\text{-}N'} \qquad q_N \simeq \frac{\eta}{2} + \left(\frac{dq}{da}\right)_{N\text{-}N} \Delta a = \frac{\eta}{2} - \eta\epsilon$$

H.O.T.

Hence, if $q_N > q_{se}$, i.e., if $\eta/2 - \cancel{\eta\epsilon} > 2\epsilon$, or $4\epsilon/\eta < 1$, point 1 moves to point 2 and indicates as unstable solution. This is in aggreement with the conclusion reached earlier via the Hamiltonian approach.

REFERENCES

1. "New Natural Satellites of the Earth?", Sky and Telescope, Vol. XXII, No. 2, July 1962, p. 10.

2. "More About the Earth's Cloud Satellites," Sky and Telescope, Vol. XXII, No. 2, August 1962, p. 63.

3. Tapley, B. D., and J. M. Lewallen, "Solar Influence on Satellite Motion Near the Stable Earth-Moon Libration Points," AIAA J. Vol. 2, No. 4, April 1964, pp. 728-732.

4. Schechter, H. B., and W. C. Hollis, Stability of the Trojan Points in the Four Body Problem, The RAND Corporation, RM-3992-PR, September 1964.

5. Feldt, W. T., and Y. Shulman, "More Results on Solar Influenced Libration Point Motion," AIAAJ., August 1966, Technical Comments - p. 1501.

6. Breakwell, J. V., and R. Pringle, Jr., "Resonances Affecting Motion Near the Earth-Moon Equilateral Libration Points," Preprint 65-683, AIAA/ION Astrodynamics Specialist Conference, Monterey, California, September 16-17, 1965; "Progress in Astronautics," Vol. 17, Academic Press, Inc., New York, 1966, pp. 55-73.

7. Smart, W. M., "Celestial Mechanics," Longmans, Green and Company, London, New York, Toronto, 1953.

8. Brown, E. W., "An Introductory Treatise on the Lunar Theory," Dover Publications, Inc., New York, 1960.

9. Moulton, F. R., "An Introduction to Celestial Mechanics," Second revised edition, tenth printing, 1958.

10. Goldstein, H., "Classical Mechanics," Addison-Wesley Publication Co., Reading, Massachusetts, London, June 1959.

11. McLachlan, N. W., "Ordinary Non-linear Differential Equations in Engineering and Physical Sciences," 2nd Edition, Clarendon Press, Oxford, 1956.